

Note: This is the final draft of the VHT roadmap submitted for the purpose of review by the project reviewers. Upon approval, the text will be transferred from the deliverable layout to the roadmap layout, including a make-over several figures to make the style coherent.



Building the European Virtual Human Twin

Call: Accelerating best use of technologies (DIGITAL-2021-DEPLOY-01)

Work program year: DIGITAL-2021-2022

Topic: ID DIGITAL-2021-DEPLOY-01-TWINS-HEALTH

Grant Agreement No: 101083771

Deliverable 3.2 VHT roadmap

Due date of delivery: 31 December 2024

Actual submission date: 14 January 2025 (deliverable was closed for submission)

Start of the project: 01 October 2022

End date: 31 December 2024

Reference

Name	EDITH_D3.1_Vision_roadmap outline_VPH_Mar2023
Lead beneficiary	Virtual Physiological Human institute (VPHi)
Author(s)	Entire EDITH consortium, final editing Liesbet Geris (VPHi)
Dissemination level	Public
Type	Report
Official delivery date	31 December 2024
Date of validation by the WP Leader	31 December 2024
Date of validation by the Coordinator	31 December 2024
Signature of the Coordinator	31 December 2024

Version log

Date	Version	Involved	Comments
31/07/2023	v1.0	Entire ecosystem	Submitted first draft
15/06/2024	v2.0	Consortium	Adding information from events and deliverables
15/11/2024	v3.0	Entire ecosystem	Open writing process. Extended summary of the roadmap with recommendations was also made available until 10/12/2024
15/12/2024	v4.0	WP leaders	New structure & final recommendations.
31/12/2024	Final	PC	Final review, formal checking by PC. Submission was not possible as the deliverable was closed in the system.
14/1/2025	Final	PC	Final submitted version.

Foreword

TBC.

*“To achieve great things, two things are needed:
a plan, and not quite enough time”*
Leonard Bernstein

Executive Summary

The Virtual Human Twin (VHT) is a groundbreaking initiative aimed at revolutionizing healthcare by creating personalized, comprehensive, and dynamic digital representations of individuals. It is envisioned as a systematic, ever-growing digital and quantitative representation of actionable knowledge in human pathophysiology, encompassing data, models, and insights from diverse disciplines. The VHT will transform healthcare by improving diagnostics, enabling personalized treatment plans, and enhancing clinical decision-making. However, realizing this ambitious vision requires addressing a myriad of technical, infrastructural, ethical, legal, and social implications.

This document presents a **comprehensive roadmap** outlining the essential elements and strategies for the successful development and implementation of the VHT. The roadmap is structured in six distinct parts, each focusing on a specific aspect crucial to the initiative's success.

Part 1 establishes the **rationale and context for the VHT**, exploring the convergence of global healthcare needs, emerging trends, and technological advancements that drive its development. This section emphasizes the importance of stakeholder engagement, involving researchers, clinicians, patients, policymakers, and industry representatives in a collaborative effort to shape the VHT's vision and ensure its relevance. It delves into the key concepts of Digital Twins (DTs), highlighting their role as fundamental building blocks for the VHT. This part also provides an in-depth analysis of stakeholder needs and challenges, identifying potential barriers to the adoption of the VHT and outlining strategies to mitigate them.

Part 2 of the roadmap focuses on the **technological foundations of the VHT**, exploring how to effectively organize and integrate diverse resources, such as data, models, workflows, services and tools, to establish a robust and interoperable platform. Advanced data-generating technologies allow to capture the intricacies of human biology and (patho)physiology. DT models are at the core of the VHT, ranging from data-driven, utilizing AI, to knowledge-driven, built upon mechanistic understanding. The credibility of resources needs to be rigorously established and documented, as it provides a quantitative measure of trustworthiness and reliability for the data and models within the VHT. Integration of resources to create multi-scale and multi-organ digital twins necessitates standardized approaches to characterize and annotate models. AI tools can be leveraged to enhance data quality, accelerate resource integration, and identify gaps in the current knowledge base of the VHT.

Part 3 delves into the crucial infrastructural elements required to support the development, deployment, and sustainable operation of the VHT. This part emphasizes the need for a robust and secure infrastructure capable of handling the complexities of data storage, model integration, and collaborative research within the VHT ecosystem. To realize the ambitious goals of the VHT, a trinity of interconnected software components is required: the **Catalogue, the Repository, and the Platform**. The Catalogue serves as the central hub for discovering and accessing VHT resources, providing detailed metadata and facilitating efficient resource retrieval. Complementing the Catalogue, the Repository provides secure and reliable storage for the vast amounts of data and models that constitute the VHT, ensuring their accessibility and preservation. The Platform, built on top of the Catalogue and Repository, offers a suite of tools and services to enable users to interact with the VHT resources, execute simulations, and analyse results. This part emphasizes the importance of adopting open standards and interoperable technologies to facilitate seamless data exchange and collaboration with other initiatives and infrastructures in the EU.

Part 4 focuses on the essential **ethical, legal, social, and regulatory considerations** that are crucial for the responsible development and implementation of the VHT. This part emphasizes the need for a comprehensive framework that addresses the complex ethical, legal, and societal implications of this transformative technology. Establishing a robust and interoperable VHT requires the adoption of common standards for data formats, model descriptions, and terminology, to ensure data quality, facilitate data integration, and enable meaningful comparisons across studies and populations. Additionally, a robust ethical and legal framework must be established to ensure the responsible use of

VHT technology and protect individual rights and privacy. This includes addressing data security, informed consent, data ownership, and the potential for bias and discrimination. Following Responsible Research and Innovation principles will be crucial in navigating these complexities, to establish trust and facilitate social acceptance of the VHT.

Part 5 delves into the crucial aspects of **engaging users, promoting the adoption of the VHT, and ensuring its long-term sustainability**. This part emphasizes the need for a multifaceted approach that incentivizes participation, addresses user concerns, and fosters a vibrant VHT ecosystem. It emphasizes the importance of creating a user-friendly interface, providing clear documentation, and offering training and support to facilitate user adoption. This part furthermore explores various business model strategies, including public funding, private investment, subscription models, and data licensing agreements, highlighting the importance of a phased approach that adapts to the evolving landscape of the VHT ecosystem. Establishing a VHT Marketplace is envisioned as a key facilitator for connecting VHT creators and consumers, promoting the exchange of resources, and fostering a sustainable economic model for the initiative. The roadmap also underscores the significance of integrating the VHT with existing research infrastructures and leveraging national and European funding opportunities to secure long-term sustainability.

Finally, part 6 brings together the **key findings and recommendations** from the previous parts, outlining a tentative timeline of activities for the successful rollout of the VHT over the next decade. It stresses the need for a collaborative effort involving all stakeholder groups to realize the full potential of this transformative technology and integrate it effectively into the healthcare landscape. This section highlights the importance of continuous investment in research and development to advance VHT technologies, expand the knowledge base, and address emerging challenges. It also emphasizes the need for ongoing dialogue and collaboration among stakeholders to ensure that the VHT evolves responsibly and equitably, ultimately contributing to the well-being of individuals and society as a whole.

The VHT roadmap provides a comprehensive framework and set of recommendations for stakeholders to effectively contribute to the development and implementation of this transformative technology. Embracing the principles outlined in this roadmap will pave the way for a future where healthcare is personalized, predictive, and participatory, leading to improved health outcomes for all.

Editors and contributors

Editors

Liesbet Geris, PhD, *VPH Institute*

Edwin Morley-Fletcher, *Lynkeus*

Amaryllis Raouzaïou, PhD, *ATHENA*

Sabato Mellone, PhD, *UNIBO*

Contributors

Note: a list of all contributors will be added as Annex 1 to this roadmap. The individuals listed as contributors to the VHT roadmap are people who participated in public EDITH-CSA meetings (on-site or on-line), expert meetings, one-on-one meetings, stakeholder engagement activities and conferences. In addition, it contains the names of individuals that contributed to the writing of the roadmap or EDITH-CSA deliverables, in EDITH-CSA activities and realizations, or contributed success stories, use cases or user stories. In total, **over 820 individuals** from all stakeholder groups in the ecosystem contributed in one way or another to the realisation of this roadmap.

Disambiguation

EDITH is a Coordination and Support Action funded by the Digital Europe program of the European Commission under the grant agreement n.101083771. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

The VHT infrastructure that is described in this roadmap, is a proposal based on the extensive and inclusive ecosystem-centred process that was followed by the EDITH consortium, as described at the start of this roadmap. A tender process has been started by the European Commission, independently from the EDITH project, to build the VHT's advanced simulation platform¹.

¹ https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/tenderdetails/docs/16cc3c6a-844a-42d4-9746-dcc7444b8001-CN/Technical%20Specifications_EC-CNECTLUX-2024-OP-0014_Advanced%20VHT%20modelling%20platform_V1.pdf.

Table of Contents

FOREWORD	3
EXECUTIVE SUMMARY	4
EDITORS AND CONTRIBUTORS	6
TABLE OF CONTENTS.....	7
LIST OF FIGURES.....	13
LIST OF TABLES	14
LIST OF BOXES.....	15
ACRONYMS.....	17
1 GENESIS AND VALIDATION OF THE ROADMAP	26
1.1 THE EDITH-CSA PROJECT	26
1.2 GENESIS OF THE VISION AND ROADMAP OUTLINE	26
1.3 ECOSYSTEM ENGAGEMENT ACTIVITIES	27
1.4 WRITING OF THE ROADMAP	27
1.5 VALIDATION OF THE ROADMAP	28
1.5.1 General validation approach: key principles.....	28
1.5.2 Validation by advisory group of stakeholders.....	28
1.5.3 From an innovation ecosystem to an innovation framework	29
2 STRUCTURE OF THE ROADMAP	30
2.1 FROM NEED STATEMENT TO SUSTAINABILITY	30
2.2 VHT ROADMAP FOR STAKEHOLDERS	30
2.3 SUCCESS STORIES, USE CASES AND USER STORIES.....	30
PART 1: FROM DIGITAL TWINS IN HEALTHCARE TO THE VIRTUAL HUMAN TWIN	31
3 DEFINITIONS.....	32
3.1 DIGITAL TWINS IN HEALTHCARE	32
3.2 DIGITAL TWINS IN HEALTHCARE – CORE COMPONENTS	32
3.2.1 Hardware Components	32
3.2.2 Middleware Components.....	33
3.2.3 Software Components.....	33
3.3 DIGITAL TWIN MODELS – KEY CONCEPTS	33
3.4 THE LIFE CYCLE OF DIGITAL TWIN MODELS.....	34
3.5 THE VIRTUAL HUMAN TWIN.....	36
4 GLOBAL TRENDS IN THE EUROPEAN CONTEXT	38
4.1 TREND 1: ESCALATING HEALTHCARE COSTS, WORKFORCE SHORTAGE AND THE NEED FOR EFFICIENCY.....	38
4.2 TREND 2: THE IMPERATIVE FOR PERSONALIZED MEDICINE	38
4.3 TREND 3: THE RISE OF DIGITALIZATION IN HEALTHCARE	39
4.4 TREND 4 : THE EU DIGITAL AGENDA AND THE EUROPEAN LIFE SCIENCES RESEARCH INFRASTRUCTURE	39
4.5 TREND 5: CONTRIBUTING TO THE SUSTAINABLE DEVELOPMENT GOALS	40
5 VISION FOR THE VIRTUAL HUMAN TWIN.....	41
5.1 THE VISION FOR THE VHT	41
5.2 THE CURRENT STATUS OF THE VHT	41
5.3 A GLANCE INTO THE FUTURE OF THE VHT	45
5.3.1 User story: patient perspective – ultimate VHT beneficiary.....	45
5.3.2 User story: clinician & healthcare provider perspective – VHT user and beneficiary.....	46
5.3.3 User story: researcher perspective – VHT user, developer and beneficiary	48
5.3.4 User story: industry perspective – VHT infrastructure accelerates drug development.....	49
6 THE VIRTUAL HUMAN TWIN ECOSYSTEM	51
6.1 VHT BENEFICIARIES AND USERS.....	52
6.1.1 Patients & Citizens	52

6.1.2	<i>Healthcare professionals & systems</i>	53
6.2	VHT CREATORS	54
6.2.1	<i>Research and innovation actors</i>	54
6.2.2	<i>Industry DT creators</i>	55
6.3	VHT ENABLERS	56
6.3.1	<i>Infrastructure actors for Data, Research and Computing</i>	56
6.3.2	<i>Technology investors, accelerators, governmental and non-governmental funding agencies</i>	56
6.4	VHT MANAGERS	56
6.4.1	<i>Policy & Law makers</i>	56
6.4.2	<i>Regulators, standardization actors, testing and clinical trial actors</i>	57
6.4.3	<i>Payers, Buyers, Reimbursement decision-makers, Health Technology Assessors</i>	59
6.4.4	<i>Legal, IP, Ethical, & Social actors</i>	61
6.5	VHT CATALYSTS	62
6.5.1	<i>Professional associations, Trade bodies, Civil society organizations</i>	62
6.5.2	<i>Education, training, communication actors</i>	62
7	NEEDS, CHALLENGES AND BARRIERS TO THE VIRTUAL HUMAN TWIN	64
7.1	ANALYSIS OF STAKEHOLDER NEEDS AND CHALLENGES	64
7.1.1	<i>Patients & Citizens</i>	64
7.1.2	<i>Clinical Community, Healthcare Providers, and Hospitals</i>	64
7.1.3	<i>Academia</i>	65
7.1.4	<i>Industry</i>	65
7.1.5	<i>Regulatory, Standardization, Legal, Ethical and Social Sciences Actors</i>	66
7.1.6	<i>Payer, Buyer, and Health Technology Assessment Community</i>	66
7.2	BARRIERS TO THE ADOPTION OF THE VHT	66
8	NEED AND VISION FOR VHT: CONCLUSIONS AND RECOMMENDATIONS	69
8.1	CONCLUSIONS	69
8.2	RECOMMENDATIONS	69
	PART 2: REALISING THE VHT - TECHNOLOGY	71
9	ORGANISATION OF RESOURCES FOR VHT	72
9.1	MULTIDIMENSIONAL SPACE AS AN ORGANISATIONAL PARADIGM	72
9.2	THE DATA OBJECT	72
9.2.1	<i>The data object type and pose</i>	72
9.2.2	<i>Annotation and annotation services</i>	73
9.2.3	<i>An illustration of the data object pose and its six dimensions</i>	76
9.3	THE MODEL OBJECT	77
9.3.1	<i>The model object type and pose</i>	77
9.3.2	<i>Execution, storage and networking services</i>	77
9.4	WORKFLOW OBJECTS	78
9.5	THE CREDIBILITY AXIS	78
9.5.1	<i>Accuracy and credibility</i>	78
9.5.2	<i>Quantifying credibility</i>	79
10	DATA AND DATA GENERATING HARDWARE FOR VHT	81
10.1	DATA GENERATION SOURCES AND SENSING TECHNOLOGIES	81
10.2	ADVANCED HARDWARE ARCHITECTURES AND HARDWARE-SOFTWARE CODESIGN	84
10.3	INTEGRATING DIVERSE DATA SOURCES	85
10.3.1	<i>Clinical grade data</i>	85
10.3.2	<i>Research grade data</i>	85
10.3.3	<i>Data generated or transformed by models</i>	86
10.4	DATA USE AND REUSE	86
10.5	DATA TRANSFORMATION SERVICES: HARMONISING AND TRANSCENDING BOUNDARIES	87
10.5.1	<i>Unit Conversion</i>	87
10.5.2	<i>Format Conversion</i>	87
10.5.3	<i>Dividing individuals in groups</i>	87
10.5.4	<i>Personal to Population Level Transitions</i>	88

10.5.5	<i>Interlinking Models: The Input-Output Perspective</i>	88
11	IN SILICO MODELS FOR VHT	89
11.1	STATE OF THE ART	89
11.1.1	<i>In silico medicine</i>	89
11.1.2	<i>Data-driven models in health and care</i>	89
11.1.3	<i>Knowledge-driven models in health and care</i>	90
11.2	MODELS AS DATA TRANSFORMATION SERVICES	92
11.3	MODELS AS DATA GENERATION SERVICES	93
11.4	MODELS AS DATA FLOW ORCHESTRATIONS	94
12	INTEGRATION OF RESOURCES	95
12.1	IDENTIFICATION OF POSSIBILITIES FOR INTEGRATION	95
12.1.1	<i>Integration of multiscale models</i>	95
12.1.2	<i>Workflows</i>	96
12.1.3	<i>AI tools for resource integration</i>	97
12.2	INTEGRATION OF RESOURCES AS A BASIS FOR KNOWLEDGE FINDING	97
12.2.1	<i>Large Language Models</i>	98
12.2.2	<i>Knowledge graphs and LLMs: A Synergistic Framework</i>	99
13	DIGITAL TWIN USE CASES	101
14	TECHNOLOGY FOR VHT: CONCLUSIONS AND RECOMMENDATIONS	106
14.1	CONCLUSIONS	106
14.2	RECOMMENDATIONS	107
	PART 3: REALISING THE VHT – INFRASTRUCTURE	108
15	INFRASTRUCTURE	109
15.1	A TRINITY OF SOFTWARE	109
15.2	STRATEGIC PILLARS	109
15.2.1	<i>Pillar 1: Distributed/federated architecture</i>	109
15.2.2	<i>Pillar 2: Governance</i>	110
15.2.3	<i>Pillar 3: Openness</i>	110
15.2.4	<i>Pillar 4: User engagement</i>	110
16	CATALOGUE AND REPOSITORY	112
16.1	CATALOGUE & REPOSITORY: CONCEPTUAL REPRESENTATION	112
16.2	STATE OF THE ART	116
16.2.1	<i>Initiatives inside the EU</i>	116
16.2.2	<i>Initiatives outside of the EU – Trusted Research Environments</i>	117
16.2.3	<i>Repositories and data management tools</i>	118
16.3	CATALOGUE AND REPOSITORY WITHIN THE VHT INFRASTRUCTURE	121
16.4	CATALOGUE & REPOSITORY REQUIREMENTS	122
16.4.1	<i>Requirements for both Catalogue and Repository</i>	122
16.4.2	<i>Requirements for the Repository</i>	123
16.4.3	<i>Additional considerations</i>	123
17	GOVERNANCE PRINCIPLES	124
17.1	ROLES AND RESPONSIBILITIES	124
17.1.1	<i>Role category: patient/citizen</i>	125
17.1.2	<i>Role category: healthcare professional</i>	125
17.1.3	<i>Role category: creator/Model developer</i>	125
17.1.4	<i>Role category: platform administrator</i>	126
17.2	DECISION-MAKING PROCESSES	126
17.2.1	<i>Populating the VHT</i>	126
17.2.2	<i>Standardization procedure</i>	128
17.3	COMPLIANCE AND LEGAL CONSIDERATIONS	130
17.3.1	<i>Premises and key issues</i>	130
17.3.2	<i>Data protection</i>	131
17.3.3	<i>Data sharing</i>	132

17.3.4	<i>Licenses for utilization of resources</i>	133
18	SIMULATION PLATFORM – THE BASIS	134
18.1	GENERAL CONSIDERATIONS	134
18.2	GENERAL PRINCIPLES	135
18.3	VHT PLATFORM USER PROFILES AND USER ROLES	136
18.4	THE TECHNOLOGICAL STARTING POINT	137
18.4.1	<i>Scientific and computational workflows and workflow engines</i>	137
18.4.2	<i>Jupyter notebooks</i>	138
18.4.3	<i>Infrastructure resource usage cost</i>	138
18.4.4	<i>Authorization and accounting</i>	139
18.4.5	<i>Trusted research environments and the ‘five Safe principles’</i>	139
18.4.6	<i>Computational infrastructure</i>	140
18.5	COLLABORATIONS WITH OTHER PLATFORMS	141
19	SIMULATION PLATFORM – THE ARCHITECTURE	144
19.1	HIGH LEVEL ARCHITECTURE	144
19.2	COMPONENTS AND TECHNOLOGIES USED IN EDITH POC INFRASTRUCTURE	146
19.3	COMPUTATIONAL ARCHITECTURE	147
19.3.1	<i>Hardware - physical layer</i>	147
19.3.2	<i>Software layer</i>	148
19.3.3	<i>Access layer / Access APIs</i>	149
19.4	TESTING AND QUALITY ASSURANCE	150
19.4.1	<i>Testing</i>	150
19.4.2	<i>Quality Assurance</i>	151
20	DIGITAL TWIN USE CASES	152
21	INFRASTRUCTURE FOR VHT: CONCLUSIONS AND RECOMMENDATIONS	156
21.1	CONCLUSIONS	156
21.2	RECOMMENDATIONS	156
PART 4: REALISING THE VHT – ELSI, STANDARDS & REGULATORY	158
22	STANDARDS FOR VHT	159
22.1	STANDARDS FOR DATA FORMATS, DATA INTEGRATION AND DATA INPUT INTO MODELS	159
22.1.1	<i>Standards for data formats, integration, interoperability and access</i>	159
22.1.2	<i>Standards for human data collection and formatting</i>	161
22.1.3	<i>Standards for clinical practice</i>	161
22.1.4	<i>Other data format types</i>	162
22.1.5	<i>Standards for the description models</i>	162
22.2	STANDARDS FOR THE MODELLING PROCESS	163
22.2.1	<i>Data preparation</i>	164
22.2.2	<i>Standards for Models</i>	164
22.3	STANDARDS FOR METADATA OF DATA AND MODELS – SEMANTIC ANNOTATION AND TAXONOMY	165
22.3.1	<i>Metadata requirements</i>	166
22.3.2	<i>Minimum reporting guidelines for Clinical Decision Support Systems (CDSS)</i>	167
22.3.3	<i>Terminologies/ontologies for description and annotation of data and model (components)</i>	168
22.3.4	<i>Clinical languages, terminologies and code systems</i>	168
22.4	SEMANTIC ANNOTATION AND TAXONOMY FOR THE VHT	169
22.5	MISSING STANDARDS STILL UNDER DEVELOPMENT	169
23	REGULATORY SCIENCE	172
23.1	GENERAL CONTEXT AND REGULATORY ACTORS	172
23.1.1	<i>The intended use of DTs</i>	172
23.1.2	<i>Regulatory actors and guidance</i>	172
23.2	CURRENT STATUS OF REGULATORY PATHWAYS FOR DT MODELS	173
23.2.1	<i>Standards as drivers for regulatory science</i>	173
23.2.2	<i>Key regulatory pathways</i>	173
23.2.3	<i>Successful examples</i>	174

23.3	DTs FOR CLINICAL DECISION SUPPORT AND PERSONAL HEALTH FORECASTING	174
23.4	DTs FOR <i>IN SILICO</i> TRIALS	175
23.5	REGULATORY SCIENCE FOR THE VHT	176
23.6	DIGITAL TWIN USE CASES	177
24	HEALTH TECHNOLOGY ASSESSMENT AND PAYERS	180
24.1	GENERAL CONSIDERATIONS	180
24.2	HTA FRAMEWORKS FOR PREDICTION OF DT TECHNOLOGIES	181
24.3	HTA FRAMEWORKS AND CATEGORIZATION OF DT FOR REIMBURSEMENT DECISIONS WITHIN EU	181
24.3.1	<i>Digital Health Technologies</i>	181
24.3.2	<i>Examples of reimbursed DTs</i>	183
24.4	RELEVANCE OF PRE-COMMERCIAL AND PUBLIC PROCUREMENT OF INNOVATIONS FOR VHT	184
25	LEGAL ASPECTS.....	185
25.1	INTRODUCTION	185
25.1.1	<i>Anonymous data and legal complexities</i>	186
25.1.2	<i>Synthetic Data</i>	187
25.1.3	<i>Towards a Harmonized Framework</i>	188
25.2	PRIVACY AND DATA PROTECTION.....	188
25.3	DATA GOVERNANCE	189
25.3.1	<i>European Health Data Space</i>	189
25.3.2	<i>Data Governance Act And Data Act</i>	191
25.4	ARTIFICIAL INTELLIGENCE	192
25.5	MEDICAL DEVICES	194
25.6	CLINICAL TRIALS.....	194
25.7	IPR MANAGEMENT	195
25.8	DIGITAL TWIN USE CASES	197
25.9	VHT-RELATED REMARKS AND RECOMMENDATIONS BASED ON CURRENT EU POLICIES	198
25.10	TOWARDS A VHT CODE OF CONDUCT	200
26	THE ETHICAL DIMENSION OF THE VHT	202
26.1	PRIVACY, ACCURACY, AND OWNERSHIP IN VHT ETHICS	202
26.1.1	<i>Privacy and data protection</i>	202
26.1.2	<i>Accuracy and misrepresentation</i>	203
26.1.3	<i>Ownership and control</i>	203
26.2	AUTONOMY, FAIRNESS, AND THE RISKS OF OVER-RELIANCE ON VHT TECHNOLOGY	203
26.2.1	<i>Patient autonomy</i>	203
26.2.2	<i>Fairness and equality</i>	204
26.2.3	<i>Managing technological over-dependence</i>	204
26.3	MULTIDISCIPLINARY APPROACH TO ADDRESS ETHICAL CHALLENGES.....	204
26.3.1	<i>The principle of identity</i>	204
26.3.2	<i>Ethics of human enhancement</i>	205
26.3.3	<i>Ethics-by-design</i>	205
27	THE SOCIAL IMPACT OF THE VHT	206
27.1	RESPONSIBLE RESEARCH AND INNOVATION	206
27.1.1	<i>From ELSI to RRI</i>	206
27.1.2	<i>Stakeholder engagement efforts</i>	207
27.2	WHAT IS A DT AND WHO USES IT?.....	208
27.2.1	<i>The social dimension of data-related concerns</i>	208
27.2.2	<i>Roles of clinicians and patients</i>	208
27.2.3	<i>Further clarifications</i>	209
27.3	THE IMPORTANCE OF EARNING TRUST FOR THE ADOPTION OF THE VHT.....	209
27.3.1	<i>Key drivers of trust</i>	209
27.3.2	<i>Literacy and awareness building</i>	210
27.3.3	<i>From trust to trustworthiness</i>	211
27.4	VHT AND CONTINUOUS STAKEHOLDER ENGAGEMENT.....	212
28	ELSI, STANDARDS AND REGULATORY FOR VHT: CONCLUSIONS AND RECOMMENDATIONS	213

28.1	CONCLUSIONS	213
28.2	RECOMMENDATIONS	214
PART 5: REALISING THE VHT – USERS, UPTAKE & SUSTAINABILITY.....		216
29	INCENTIVES FOR UPTAKE OF THE VHT	217
29.1	USERS OF THE VHT	217
29.1.1	<i>Uptake of the VHT by users: general concepts</i>	<i>217</i>
29.2	INCENTIVES FOR ACADEMIC AND INDUSTRIAL UPTAKE	217
29.2.1	<i>Stakeholder engagement to identify incentives</i>	<i>217</i>
29.2.2	<i>Incentives for VHT developers</i>	<i>218</i>
29.3	CLINICAL UPTAKE AND ENGAGEMENT WITH THE VHT	219
29.3.1	<i>Stakeholder engagement to identify incentives</i>	<i>219</i>
29.3.2	<i>Incentives for VHT consumers</i>	<i>219</i>
29.4	FACILITATING ACCESS TO VHT TECHNOLOGIES THROUGH AR/VR	222
29.4.1	<i>VR & VHT to enhance training: immersive learning.</i>	<i>222</i>
29.4.2	<i>VR & VHT to support procedural planning in complex medical areas.</i>	<i>222</i>
29.4.3	<i>VR & VHT for improving communication with patients and their empowerment.</i>	<i>223</i>
29.4.4	<i>VR & VHT: an emerging opportunity for business models.</i>	<i>223</i>
30	BUSINESS MODELS FOR VHT.....	226
30.1	CURRENT & FUTURE MARKET APPRAISAL.....	226
30.1.1	<i>Growth across all applications</i>	<i>226</i>
30.1.2	<i>Growth across all End Users</i>	<i>227</i>
30.1.3	<i>Growth across all geographical regions</i>	<i>228</i>
30.1.4	<i>Conclusion of the Landscape Analysis</i>	<i>230</i>
30.2	BUSINESS MODELS FOR VHT	230
30.2.1	<i>Business model strategies and approaches.....</i>	<i>230</i>
30.2.2	<i>Trends & risks</i>	<i>233</i>
30.2.3	<i>A phased approach.....</i>	<i>233</i>
30.3	DIGITAL TWIN USE CASES.....	236
31	TOWARDS A VHT MARKETPLACE.....	239
31.1	ADOPTION PHASES.....	239
31.1.1	<i>Honour-ledger phase.....</i>	<i>239</i>
31.1.2	<i>Token-ledger phase</i>	<i>240</i>
31.1.3	<i>Market phase</i>	<i>241</i>
31.1.4	<i>Stakeholder Engagement Across Ecosystem Phases</i>	<i>241</i>
31.2	DISTRIBUTED LEDGER.....	242
31.2.1	<i>Blockchain networks.....</i>	<i>242</i>
31.2.2	<i>Smart Contracts.....</i>	<i>243</i>
31.3	EQUITABLE RESOURCE PRICING.....	243
31.3.1	<i>A game-theoretic framework</i>	<i>244</i>
31.4	A USE CASE & THE MARKETPLACE ARCHITECTURE.....	244
32	EUROPEAN PUBLIC INFRASTRUCTURE	247
32.1	STATE OF THE ART EUROPEAN INFRASTRUCTURES	247
32.2	ERICs AND EDICs.....	248
32.2.1	<i>European Research Infrastructure Consortium</i>	<i>248</i>
32.2.2	<i>European Digital Infrastructure Consortium</i>	<i>249</i>
32.3	FUNDING PROGRAMS RELEVANT FOR VHT	249
32.3.1	<i>European programs.....</i>	<i>249</i>
32.3.2	<i>Member state programs: some examples.....</i>	<i>251</i>
32.3.3	<i>Public-private partnerships for VHT sustainability</i>	<i>253</i>
32.4	TOWARDS AN ESTABLISHED VHT PUBLIC INFRASTRUCTURE	254
33	USERS, UPTAKE AND SUSTAINABILITY FOR VHT: CONCLUSIONS AND RECOMMENDATIONS.....	256
33.1	CONCLUSIONS	256
33.2	RECOMMENDATIONS	256
33.2.1	<i>Users & inclusiveness.....</i>	<i>256</i>

33.2.2	<i>Sustainability</i>	258
PART 6:	RECOMMENDATIONS FOR A SUCCESSFUL VHT	259
34	RECOMMENDATIONS FOR THE ROLL-OUT OF THE VIRTUAL HUMAN TWIN	260
34.1	A TENTATIVE TIMELINE OF ACTIVITIES	260
34.2	NEED ASSESSMENT OF CREATORS AND CONSUMERS	262
34.3	VHT TECHNOLOGIES: BASIC BUILDING BLOCKS.....	262
34.4	VHT INFRASTRUCTURE	263
34.5	ELSI, STANDARDS & REGULATORY	264
34.6	USERS & INCLUSIVENESS.....	265
34.7	SUSTAINABILITY	266
35	RECOMMENDATIONS FOR THE STAKEHOLDERS OF THE VHT ECOSYSTEM	268
35.1	EUROPEAN COMMISSION.....	268
35.2	EU MEMBER STATES	268
35.3	RESEARCH COMMUNITY.....	268
35.4	INDUSTRY	269
35.5	HEALTHCARE PROVIDERS	269
35.6	REGULATORY BODIES AND STANDARDS ORGANISATIONS.....	269
35.7	PAYER, BUYER, AND HEALTH TECHNOLOGY ASSESSMENT COMMUNITY	270
35.8	ETHICAL, LEGAL, AND SOCIAL IMPLICATION EXPERTS	270
35.9	PATIENTS AND THE PUBLIC	270
36	ANNEX 1: CONTRIBUTORS	271
37	ANNEX 2: STANDARDS AND STANDARD DEFINING ORGANISATIONS	289

List of Figures

FIGURE 1:	KEY COMPONENTS OF DIGITAL TWINS.....	32
FIGURE 2:	OVERVIEW OF EXISTING DT SOLUTIONS. THIS IS A PLACEHOLDER. THE INFORMATION IN THE BOXES WILL CONTAIN A 1-SENTENCE EXPLANATION OF THE TWIN, ITS INTENDED USE, USER, STATUS (RESEARCH/CLINICS/COMPANY)	42
FIGURE 3:	SCHEMATIC DEPICTION OF THE VHT COMMUNITY OF PRACTICE (COP) AND ITS DIFFERENT STAKEHOLDERS. ELSI: ETHICAL, LEGAL, SOCIAL ISSUES; CSO: CIVIL SOCIETY ORGANISATION.	51
FIGURE 4:	INTERDISCIPLINARITY REQUIRED FOR BUILDING IMMUNE DIGITAL TWINS	63
FIGURE 5:	A GRAPHIC REPRESENTATION OF THE DATA SHEET DESCRIBING THE RELEVANT INFORMATION OF A DATA OBJECT.....	73
FIGURE 6:	THE DIGITAL TWIN CONCEPT FOR PERSONALISED MEDICINE. A. AN INDIVIDUAL PATIENT HAS A LOCAL SIGN OF DISEASE (RED). B. A DIGITAL TWIN OF THIS PATIENT IS CONSTRUCTED IN UNLIMITED COPIES, BASED ON COMPUTATIONAL NETWORK MODELS OF THOUSANDS OF DISEASE-RELEVANT VARIABLES. C. EACH TWIN IS COMPUTATIONALLY TREATED WITH ONE OR MORE OF THE THOUSANDS OF DRUGS. THIS RESULTS IN DIGITAL CURE OF ONE PATIENT (GREEN). D. THE DRUG THAT HAS THE BEST EFFECT ON THE DIGITAL TWIN IS SELECTED FOR TREATMENT OF THE PATIENT	83
FIGURE 7:	A VIRTUAL LIVER SAMPLE TWIN. (A) LIVER IS COMPOSED OF THOUSANDS (MOUSE) TO MILLIONS (HUMAN) REPETITIVE ANATOMICAL AND FUNCTIONAL UNITS, CALLED LOBULES, THAT CAN STATISTICALLY BE APPROXIMATED BY HEXAGONS. WITHIN EACH LOBULE, METABOLISM OF AMMONIA IN ZONATED I.E., DIFFERENT REACTIONS ARE EXECUTED IN THE PERI-PORTAL, MID-ZONAL AND PERI-CENTRAL REGION (INDICATED BY THE ITALIC NUMBERS). THE PRECISE LOBULAR MICRO-ARCHITECTURE HAS BEEN COMPUTED FROM CONFOCAL MICROGRAPHS (B) A SET OF REACTIONS IN EACH ZONE OF A TISSUE SAMPLE (C) CAN BE SOLVED TO SIMULATE THE DETOXIFICATION OF BLOOD FROM AMMONIA (COLOURS).	91
FIGURE 8:	A COMPUTER MODEL AS DATA TRANSFORMATION SERVICE.....	93
FIGURE 9:	EXAMPLE OF WORKFLOW GRAPH	94
FIGURE 10:	INTEGRATION OF PATIENT'S SEQUENCING DATA AND MOLECULAR MECHANISMS IN A CELL-BASED SIMULATION FRAMEWORK TO CAPTURE INDIVIDUALS' GENETIC AND ENVIRONMENTAL INFLUENCES IN ORDER TO PROVIDE A DEEPER UNDERSTANDING OF THE BIOLOGICAL CONTEXT AND RESPONSE TO THERAPIES.....	96
FIGURE 11:	BLOOD GLUCOSE MEASUREMENTS (TOP), INSULIN RATES (ANTEPENULTIMATE), AND NUTRITION RATES (BOTTOM) ARE USED TO IDENTIFY PATIENT-SPECIFIC INSULIN SENSITIVITY EVOLUTION OVER TIME (SECOND). INSULIN SENSITIVITY IS CONSIDERED CONSTANT OVER 1-HOUR PERIODS. INSULIN AND NUTRITION ARE MITIGATED TO STABILIZE BLOOD GLUCOSE IN THE CLINICALLY-SPECIFIED SAFE TARGET BAND (GREEN AREA).....	102

FIGURE 12: HIGH-LEVEL WORKFLOW OF THE BBCT SOLUTION	102
FIGURE 13: UISS-TB LAYERS AND THEIR IMPLEMENTATION, AIMED TO SIMULATE MYCOBACTERIUM TUBERCULOSIS (MTB) DYNAMICS AND ITS INTERACTION WITH THE HUMAN IMMUNE SYSTEM ALONG WITH THE THERAPEUTIC INTERVENTIONS.	103
FIGURE 14: ATRIALMTK MODEL PIPELINE.....	104
FIGURE 15: SCHEMATIC OVERVIEW OF THE MAIN COMPONENTS OF THE VHT INFRASTRUCTURE.	109
FIGURE 16: SCHEMATIC REPRESENTATION OF THE REPOSITORY AND THE CATALOGUE DEPICTING ITS ESSENTIAL COMPONENTS.	112
FIGURE 17: HIGH-LEVEL OVERVIEW OF THE PROPOSED VHT INFRASTRUCTURE, INCLUDING THE CATALOGUE & REPOSITORY.	121
FIGURE 18: PROCEDURE TO POPULATE THE VHT	127
FIGURE 19: STANDARDIZATION PIPELINE	129
FIGURE 20: SERVICE BLUEPRINT FOR DATA OBJECT STANDARDIZATION	130
FIGURE 21: SCHEMATIC REPRESENTATION OF THE SIMULATION PLATFORM INCLUDING THE DIFFERENT LAYERS AND A (NON-EXHAUSTIVE) LIST OF SERVICES.	135
FIGURE 22: HIGH-LEVEL OVERVIEW OF THE PROPOSED VHT PLATFORM ARCHITECTURE.	144
FIGURE 23: COMPONENTS OF MVP#1	146
FIGURE 24: DETAILED VIEW OF THE USER FACING COMPONENTS OF THE EDITH PLATFORM.....	147
FIGURE 25: ARCHITECTURE OF THE MEE DEMONSTRATOR.	150
FIGURE 26: THIS FIGURE DEPICTS DATA WORKFLOW FOR THE COMPLETE GC SOLUTION TO BE COMMERCIALIZED BY INSILICARE. IN THE CONTEXT OF EDITH, THE WORKFLOW STOPS AT THE ANONYMISED DIGITAL PATIENT STEP (DASHED BLUE LINES).	152
FIGURE 27: HIGH-LEVEL WORKFLOW OF THE BBCT SOLUTION	153
FIGURE 28: TOP: OVERALL WORKFLOW FOR PERSYST ESI POWERED BY EPILOG. BOTTOM: DETAILS OF THE CLINICIAN'S INTERFACE.	154
FIGURE 29: ATRIALMTK DATAFLOW, FORMATS & COMPUTATIONAL REQUIREMENTS.	155
FIGURE 30: SEMANTIC ANNOTATION. ADAPTED FROM.....	166
FIGURE 31: BOSTON CHILDREN'S HOSPITAL 3D MODELLING WORKFLOW	184
FIGURE 32: OVERVIEW OF THE LEGISLATION RELEVANT FOR THE VHT.	186
FIGURE 33: THE VPHI INFOKIT, INCLUDING TOOLS AND RESOURCES FOR STAKEHOLDER ENGAGEMENT	211
FIGURE 34: DIGITAL TWIN OF LIVER TISSUE AND VESSELS.....	221
FIGURE 35: DIGITAL TWINS IN THE HEALTHCARE MARKET, BY APPLICATION, 2023 vs 2028 (USD MILLION).	227
FIGURE 36: DIGITAL TWINS IN HEALTHCARE MARKET, BY END-USER, 2023 vs 2028 (USD MILLION).	228
FIGURE 37: GEOGRAPHICAL SNAPSHOT OF DIGITAL TWINS IN THE HEALTHCARE MARKET.....	228
FIGURE 38: PHASE 1 (SHORT TERM) DEVELOPMENT OF BUSINESS MODELS.	234
FIGURE 39: PHASE 2 (MIDDLE LONG TERM) DEVELOPMENT OF BUSINESS MODELS.	235
FIGURE 40: FIGURE X: PHASE 3 (LONG TERM) DEVELOPMENT OF BUSINESS MODELS.....	236
FIGURE 41: THE THREE PHASES OF THE VHT MARKETPLACE ADOPTION.	239
FIGURE 42: VHT MARKETPLACE USE CASE.	245
FIGURE 43: VHT MARKETPLACE ARCHITECTURE.	246
FIGURE 44: SCHEMATIC REPRESENTATION OF THE 4 DATA, COMPUTING AND DIGITAL RESEARCH INFRASTRUCTURES (LEFT, GREY) AND THE 16 RIS IN THE HEALTH AND FOOD CATEGORY (RIGHT, ORANGE) AS DEPICTED IN THE 2021 ROADMAP OF EUROPEAN STRATEGY FORUM ON RESEARCH INFRASTRUCTURES.	248
FIGURE 45: OVERVIEW OF THE LIFE SCIENCES RESEARCH INFRASTRUCTURES AND THE VARIOUS PROJECTS THROUGH WHICH THEY COLLABORATE TO INCREASE IMPACT	249
FIGURE 46: SCHEMATIC REPRESENTATION OF KEY ACTIVITIES FOR VHT OVER TIME	261

List of Tables

TABLE 1: EDITH INFRASTRUCTURE: A TRINITY OF SOFTWARE	121
TABLE 2: MAIN CHARACTERISTICS OF EDITH USE CASES.....	134
TABLE 3: BIOLOGICAL, TECHNICAL, LEGAL AND DESCRIPTIVE METADATA DESCRIBING A DT	170
TABLE 4: A NON-EXHAUSTIVE OVERVIEW OF SEVERAL DIGITAL HEALTH TECHNOLOGY ASSESSMENT FRAMEWORKS IN MEMBER STATES, WHICH RELATE TO OR IMPACT THE DIGITAL TWINS IN HEALTHCARE, AS ENVISIONED WITHIN THE VHT INITIATIVE.	182
TABLE 5: OVERVIEW OF IDENTIFIED SOCIAL IMPLICATIONS	207
TABLE 6: TECHNOLOGY DEVELOPMENT AND SYSTEMS INTEGRATION BUSINESSES AND THE RELATIVE TIMING TO OPERATE: 1 (SHORT TERM, 1-4 YEARS), 2 (MIDDLE LONG TERM, 5-7 YEARS) OR 3 (LONG TERM, 8-10 YEARS).	231
TABLE 7: DATA MANAGEMENT AND SECURITY BUSINESSES AND THE RELATIVE TIMING TO OPERATE: 1 (SHORT TERM, 1-4 YEARS), 2 (MIDDLE LONG TERM, 5-7 YEARS) OR 3 (LONG TERM, 8-10 YEARS).	231

TABLE 8: R&D AND CLINICAL DECISION SUPPORT BUSINESSES AND THE RELATIVE TIMING TO OPERATE: 1 (SHORT TERM, 1-4 YEARS), 2 (MIDDLE LONG TERM, 5-7 YEARS) OR 3 (LONG TERM, 8-10 YEARS)	232
TABLE 9: BUSINESS INTEGRATION AND SUPPORT SUPPORT BUSINESSES AND THE RELATIVE TIMING TO OPERATE: 1 (SHORT TERM, 1-4 YEARS), 2 (MIDDLE LONG TERM, 5-7 YEARS) OR 3 (LONG TERM, 8-10 YEARS).....	232
TABLE 10: STAKEHOLDER ENGAGEMENT ACROSS ECOSYSTEM PHASES.....	241
TABLE 11: FUNDING OPPORTUNITIES FOR THE VHT	252
TABLE A 12: TECHNICAL STANDARD DEFINING ORGANISATIONS	289
TABLE A 13: CLINICAL STANDARD DEFINING ORGANISATIONS	290
TABLE A 14: STANDARDS FOR MEDICAL IMAGING	291
TABLE A 15: STANDARD FORMATS FOR ELECTRO- AND NEUROPHYSIOLOGY, BIOSIGNAL AND VITAL SIGN DATA.....	292
TABLE A 16: STANDARDS FOR GENETIC SEQUENCE VARIANTS	293
TABLE A 17: STANDARDS FOR MODELS	294
TABLE A 18: STANDARDS FOR MODEL SIMULATIONS AND DOCUMENTATION OF RESULTS	296
TABLE A 19: TERMINOLOGIES AND ONTOLOGIES FOR THE DESCRIPTION AND ANNOTATION OF DATA, MODELS, AND THEIR COMPONENTS	297
TABLE A 20: CLINICAL LANGUAGES, TERMINOLOGIES, AND CODE SYSTEMS.....	299
TABLE A 21: GENERAL METADATA STANDARDS, FORMATS, AND PROTOCOLS	300

List of Boxes

BOX 1: SUCCESS STORY - EXAMPLE	30
BOX 2: USE CASE - GLYCEMIC CONTROL IN ICU PATIENTS (CONTEXT)	42
BOX 3: USE CASE - OSTEOPOROTIC FRACTURE RISK PREDICTION (CONTEXT)	43
BOX 4: USE CASE - UNIVERSAL IMMUNE SYSTEM SIMULATOR FOR TUBERCULOSIS (CONTEXT).....	43
BOX 5: USE CASE - EPILEPTOGENIC ZONE LOCALISATION FOR SURGICAL PLANNING IN EPILEPSY PATIENTS (CONTEXT)	44
BOX 6: USE CASE - THE ATRIAL MODELLING TOOLKIT FOR CARDIOVASCULAR DIGITAL TWINS (CONTEXT).....	44
BOX 7: USER STORY – A PATIENT’S JOURNEY WITH DIGITAL TWINS IN HEALTHCARE	45
BOX 8: USER STORY – A CLINICIAN’S PERSPECTIVE	46
BOX 9: USER STORY – A RESEARCHER’S PERSPECTIVE.	48
BOX 10: USER STORY – A COMPANY’S PERSPECTIVE.	49
BOX 11: SUCCESS STORY –THE IMMUNE DIGITAL TWIN INTIATIVE.....	63
BOX 12: SUCCESS STORY – scDrugPRIO	83
BOX 13: SUCCESS STORY – DIGIPREDICT PROJECT.....	84
BOX 14: SUCCESS STORY – RADIOMICS DIGITAL TWINS USING MACHINE LEARNING FOR CANCER DIAGNOSIS.....	90
BOX 15: SUCCESS STORY – USING DIGITAL TWINS TO ADVANCE MECHANISTIC INSIGHTS.	91
BOX 16: SUCCESS STORY – VIRTUAL TWINS FOR CANCER	95
BOX 17: SUCCESS STORY –A MULTI-LEVEL ATHEROSCLEROTIC PLAQUE GROWTH MODEL FOR CORONARY ARTERIES	96
BOX 18: EDITH DEVELOPMENT – EDITH-CSA KNOWLEDGE SOURCING AND KNOWLEDGE GRAPHS.....	100
BOX 19: USE CASE - GLYCEMIC CONTROL IN ICU PATIENTS (DIGITAL TWIN)	101
BOX 20: USE CASE - OSTEOPOROTIC FRACTURE RISK PREDICTION (DIGITAL TWIN)	102
BOX 21: USE CASE - UNIVERSAL IMMUNE SYSTEM SIMULATOR FOR TUBERCULOSIS (DIGITAL TWIN).....	103
BOX 22: USE CASE - EPILEPTOGENIC ZONE LOCALISATION FOR SURGICAL PLANNING IN EPILEPSY PATIENTS (DIGITAL TWIN)	104
BOX 23: USE CASE - THE ATRIAL MODELLING TOOLKIT FOR CARDIOVASCULAR DIGITAL TWINS (DIGITAL TWIN).....	104
BOX 24: SUCCESS STORY – THE 12 LABOURS PROJECT.....	111
BOX 25: EDITH DEVELOPMENT – MOCK-UP FOR UPLOADING A RESOURCE IN THE REPOSITORY	113
BOX 26: EDITH DEVELOPMENT – MOCK-UP FOR PREPARING A MODEL IN THE REPOSITORY	115
BOX 27: SUCCESS STORY – BUILDING THE HUMAN CELL ATLAS.....	118
BOX 28: SUCCESS STORY – DIGITAL TWINS OF THE BRAIN	119
BOX 29: SUCCESS STORY – DATA COLLECTION INITIATIVE IN GERMANY.....	120
BOX 30: EDITH DEVELOPMENT – PRE-SELECTED USE CASES INFORMING VISION AND ROADMAP	134
BOX 31: SUCCESS STORY – SEEK PLATFORM FOR HETEROGENEOUS SCIENTIFIC RESEARCH OUTPUTS.....	143
BOX 32: EDITH DEVELOPMENT – PROOF OF CONCEPT INFRASTRUCTURE.....	146
BOX 33: EDITH DEVELOPMENT – MODEL EXECUTION ENVIRONMENT	149
BOX 34: USE CASE - GLYCEMIC CONTROL IN ICU PATIENTS (WORKFLOW)	152

Box 35: USE CASE - OSTEOPOROTIC FRACTURE RISK PREDICTION (WORKFLOW)	152
Box 36: USE CASE - UNIVERSAL IMMUNE SYSTEM SIMULATOR FOR TUBERCULOSIS (WORKFLOW)	153
Box 37: USE CASE - EPILEPTOGENIC ZONE LOCALISATION FOR SURGICAL PLANNING IN EPILEPSY PATIENTS (WORKFLOW)	153
Box 38: USE CASE - THE ATRIAL MODELLING TOOLKIT FOR CARDIOVASCULAR DIGITAL TWINS (WORKFLOW)	154
Box 39: USE CASE - GLYCEMIC CONTROL IN ICU PATIENTS (REGULATORY)	177
Box 40: USE CASE - OSTEOPOROTIC FRACTURE RISK PREDICTION (REGULATORY)	177
Box 41: USE CASE - UNIVERSAL IMMUNE SYSTEM SIMULATOR FOR TUBERCULOSIS (REGULATORY)	177
Box 42: USE CASE - EPILEPTOGENIC ZONE LOCALISATION FOR SURGICAL PLANNING IN EPILEPSY PATIENTS (REGULATORY)	178
Box 43: USE CASE - THE ATRIAL MODELLING TOOLKIT FOR CARDIOVASCULAR DIGITAL TWINS (REGULATORY)	178
Box 44: SUCCESS STORY – REIMBURSEMENT OF DIGITAL TWINS FOR PRECISION PAEDIATRIC HEART SURGERY	183
Box 45: USE CASE - GLYCEMIC CONTROL IN ICU PATIENTS (LEGAL)	197
Box 46: USE CASE - OSTEOPOROTIC FRACTURE RISK PREDICTION (LEGAL)	197
Box 47: USE CASE - EPILEPTOGENIC ZONE LOCALISATION FOR SURGICAL PLANNING IN EPILEPSY PATIENTS (LEGAL)	197
Box 48: USE CASE - THE ATRIAL MODELLING TOOLKIT FOR CARDIOVASCULAR DIGITAL TWINS (LEGAL)	198
Box 49: SUCCESS STORY – THE VPHI INFO KIT FOR STAKEHOLDER ENGAGEMENT	210
Box 50: EDITH DEVELOPMENT – THE VHT INDUSTRY VALUE PROPOSITION	218
Box 51: SUCCESS STORY – THE VIRTUOUS CYCLE OF A PARTNERSHIP BETWEEN DIGITAL TWIN CREATORS AND CLINICIANS	220
Box 52: SUCCESS STORY – ICU DIGITAL TWINS, FROM RESEARCH TO CLINICAL APPLICATION	221
Box 53: SUCCESS STORY – VIRTUAL REALITY & DIGITAL TWINS IN A PAEDIATRIC CLINICAL CENTRE	223
Box 54: EDITH DEVELOPMENT – BizMOD4DTH PORTFOLIO OF BUSINESS MODELS FOR VHT	230
Box 55: USE CASE - GLYCEMIC CONTROL IN ICU PATIENTS (MARKET)	236
Box 56: USE CASE - OSTEOPOROTIC FRACTURE RISK PREDICTION (MARKET)	237
Box 57: USE CASE - UNIVERSAL IMMUNE SYSTEM SIMULATOR FOR TUBERCULOSIS (MARKET)	237
Box 58: USE CASE - EPILEPTOGENIC ZONE LOCALISATION FOR SURGICAL PLANNING IN EPILEPSY PATIENTS (MARKET)	237
Box 59: USE CASE - THE ATRIAL MODELLING TOOLKIT FOR CARDIOVASCULAR DIGITAL TWINS (MARKET)	238

Acronyms

Acronym	Full name
1+MG	1+ Million Genomes
AAI	Authentication and Authorization Infrastructure
ABE	Attribute-Based Encryption
ABM	Agent-based model
aBMD	Areal Bone Mineral Density
ACC CYFRONET	Akademickie Centrum Komputerowe Cyfronet AGH
ACP	Algorithm Change Protocol
ADR	Adverse Drug Reaction
AEs	Auto-Encoders
AFib	Atrial Fibrillation
AGS	Advisory Group of Stakeholders
AHA	American Hospital Association
AHP	American Hospital of Paris
AI	Artificial Intelligence
AI Act	Artificial Intelligence Act (Regulation (EU) 2024/1689)
AI HLEG	Artificial Intelligence High-Level Expert Group
AIB	Artificial Intelligence Board
AIM	International Association of Mutual Benefit Societies
ALCOA	Attributable, Legible, Contemporaneous, Original, Accurate
AMICE	Association of Mutual and Cooperative Insurer in Europe
APDL	Ansys Parametric Design Language
API	Application Programming Interface
AQuAS	Agència de Qualitat i Avaluació Sanitàries de Catalunya
AR	Augmented Reality
ARF0	Absolute Risk Of Current Fracture
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATHENA	Athina-Erevnitiko Kentro Kainotomias Stis Technologies Tis Pliroforias, Ton Epikoinonion Kai Tis Gnosis
ATMP	Advanced Therapy Medicinal Product
Atrialmtk	Atrial Modelling Tool Kit
AV	Anti-Virus
AWS	Amazon Web Services
B1MG	Beyond 1 Million Genomes
B2B	Business-to-Business
BAM	Binary Alignment Map
BBCT	Bologna Biomechanical CT
BBCT	Bologna Biomechanical Computed Tomography
BCO	BioCompute objects
BCR	Binding Corporate Rules
BDML	Biological Dynamics Markup Language
BIDS	Brain Imaging Data Structure
BioPAX	Biological Pathways Exchange
BMC	Business Model Canvas
BMD	Bone Mineral Density
BND	Boundary format
BRCAPRO	BRCA carrier prediction model
BSC	Barcelona Supercomputer Centre
CAGR	Compound Annual Growth Rate
CAMD	Competent Authorities for Medical Devices
CBER	Centre for Biologics Evaluation and Research
CC-BY-SA	Creative Commons Attribution Share-Alike license
CDA	Clinical Document Architecture
CDC	Center for Disease Control

CDISC	Clinical Data Interchange Standards Consortium
CDM	Clinical data management
CDMO	Contract Development and Manufacturing Organization
CDRH	Centre for Devices and Radiological Health
CE	Conformité Européenne
CellML	Markup language for mathematical models
CEN	European Committee for Standardization
CENELEC	European Electrotechnical Committee for Standardization
CERIF	Common European Research Information Format
CERN	European Organization for Nuclear Research
CFD	Computational Fluid Dynamics
CFG	Configuration
CGM	Continuous Glucose Monitor
CHEMFET	Chemical field-effect transistor
CJEU	Court of Justice of the European Union
CKAN	Comprehensive Knowledge Archive Network
CKM	Clinical Knowledge Manager
CM&S	Computer Modelling & Simulation
cMDX	Clinical Map Document based on XML
CMS	Centers for Medicare & Medicaid Services
CO	Cell Ontology
COCIR	European Trade Association representing the medical imaging, radiotherapy, health ICT and electromedical industries
COMBINE	Computational Modeling in Biology Network
CoP	Community of Practice
CoU	Context of Use
CP	Central Processor
CPM	Common Provenance Model
CPU	Central Processing Unit
CRAM	Compresses Reference-oriented Alignment Map
CRO	Contract Research Organisation
CSA	Coordination and Support Action
CSO	Civil Society Organisations
CSV	Comma-separated values
CT	Computed Tomography
CTCA	Computed Tomographic coronary angiography
CTIS	Clinical Trials Information System
CTR	Clinical Trial Regulation (Regulation (EU) 536/2014)
CVD	Cardiovascular Disease
CWL	Common workflow language
D	Deliverable
DA	Data Act Regulation (EU) 2023/2854
DAT	Data file format
DCAT-AP	Data Catalogue vocabulary – Application Profile
DCJSC	Data Contract JSON Serializer
DGA	Data Governance Act (Regulation (EU) 2022/868)
DICOM	Digital imaging and communications in medicine
DIGITALEUROPE	DIGITALEUROPE AISBL
DIKW	Data-Information-Knowledge-Wisdom
DIN	Deutschen Institut für Normung e. V.
DL	Deep learning
DLT	Distributed Ledger Technology
DMD	Digital Medical Device
DNA	Deoxyribonucleic Acid
doi	Digital object identifier
DOID	Human Disease Ontology
DOP	Data Object Pose
DOT	Data Object Type

DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
DSM	Digital Single Market
DSMS	Data Storage Management Services
DT	Digital Twin
DTD	Digital Twin Definition Language
DTH	Digital Twin in Healthcare
DTIac	Digital Twin Industry market, applications combined
DTM	Deep Thinkers Meeting
DXA	Dual-energy X-ray absorptiometry
EAB	External Advisory Board
EAP	European Association for Psychotherapy
EATRIS	European infrastructure for translational medicine
EBI	European Bioinformatics Institute
EC	European Community
ECDC	European Centre for Disease Prevention and Control
ECG	Electrocardiogram
ECRIN	European Clinical Research Infrastructure Network
EDIC	European Digital Innovation Hub
EDIH	European Digital Innovation Hubs
EDIHTA	European Digital Health Technology Assessment framework
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EFN	European Federation of Nurses Associations
EFORT	European Federation of National Associations of Orthopaedics and Traumatology
EFPIA	European Federation from Pharmaceutical industries and associations
EFSS	Enterprise File Sync and Sharing
EGC	European General Court
EGE	European Group on Ethics in Science and New Technologies
EHDS	European Health Data Space
EHR	Electronic healthcare record
EHRxF	Electronic Health Records Exchange Format
EHTEL	European Health Telematics Association
EIF	EOSC Interoperability Framework
ELIXIR	European life sciences infrastructure
ELPA	European Liver Patient Organisation
ELSI	Ethical, Legal & Social Issues
EMA	European Medicine Agency
EMBS	Engineering in Medicine & Biology Society
EMEWS	Extreme-scale Model Exploration with Swift
EMPRICA	Gesellschaft für Kommunikations- und Technologieforschung mbH
EMR	Electronic Medical Record
EMSP	European Multiple Sclerosis Platform
ENC	European Nursing Council
ENISA	European Union Agency for Cybersecurity
EOSC	European Open Science Cloud
EPC	European Patent Convention
EPCC	Edinburgh Parallel Computing Centre
EPF	European Patients' Forum
EPFL	École Polytechnique Fédérale de Lausanne
EPO	European Patent Office
ERIC	European Research Infrastructure Consortium
ESC	European Society of Cardiology
ESFRI	European Strategy Forum on Research Infrastructures
EU	European Union
EUCAIM	European Cancer Images project
EUDAMED	European Database for Medical Devices
EUDAT	European Data e-Infrastructure Initiative

EUDAT-CPI	EUDAT Collaborative Data Infrastructure
EUHA	European Hospital and Healthcare Federation
EUnetHTA	European Network for Health Technology Assessment
EUPATI	European Patients' Academy on Therapeutic Innovation
EUREC	European Network of Research Ethics Committees
EuroHPC	European High-Performance Computing
EUROoCS	European Organ-on-Chip Society
EUSIP	European Social Insurance Platform
FAIR	Findable, Accessible, Interoperable, Reusable
FASTQ	sequencing format with quality information
FDA	Food and Drug Administration
FEM	Finite Element Modelling
FFR	Fractional Flow Reserve
FHIR	Fast Healthcare Interoperability Resources
FieldML	Meta-language for representing hierarchical models using mathematical fields
FIPAT	Computational Modeling in Biology Network
FL	Federated Learning
FMA	Functional Model of Anatomy
FORTH	Foundation for Research and Technology Hellas)
FRAND	Fair, Reasonable, and Non-Discriminatory
FUTURE-AI	Fairness, Universality, Traceability, Usability, Robustness & Explainability in AI
FVM	Finite volume modelling
FZJ	Forschungszentrum Jülich GmbH
GA	Grant Agreement
GA4GH	Global Alliance for Genomics and Health
GANS	Generative Adversarial Networks
GB	Gigabyte
GBA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
GDC	Genomic Data Commons
GDI	European Genomic Data Infrastructure
GDPR	General Data Protection Regulation (Regulation (EU) 2016/679)
GLP	Good Laboratory Practice
GMLP	Good Machine Learning Practice
GMP	Good Manufacturing Practice
GO	Gene Ontology
GPL	Global public licence
GPU	Graphics Processing Unit
GSP	Goos simulation practice
GUI	Graphical User Interface
GxP	Good x Practice
HAS	Haute Agence de Santé (France)
HBP	Human Brain Project
HCA	Human Cell Atlas
HCP	Health Care Providers
hDAAs	harmonised Data Access Agreements
HDAB	Health Data Access Body
HDF	Hierarchical Data Format
HDR	Health Data Research
HE	Homomorphic encryption
HER	Health Electronic Record
HGP	Human Genome Project
HIPAA	Health Insurance Portability and Accountability Act
HITS	Heidelberg Institute for Theoretical Studies gGmbH
HL7	Health Level Seven International
HL7-STU	Health Level Seven International – Standards for Trial Use
HMAC	Hash-based Message Authentication Code
HOPE	European Hospital and Healthcare Federation

HPC	High-Performance Computing
HQ	Head Quarter
HRAI	High-Risk AI
HTA	Health Technology Assessment
HTA-CG	Health Technology Assessment Coordination Group
HTAi	Health Technology Assessment International
HTC	High-Throughput Computing
HTML	HyperText Markup Language
HTS	High-throughput screening
IaaS	Infrastructure as a Service
IAB	Industry Advisory Board
IAM	Identity & Access Management
IB	Intermediary Business
ICD	International Classification of Diseases
ICEM	Ansys mesh generator
ICING	Intensive Control Insulin-Nutrition-Glucose
ICO	Information Commissioner's Office (UK)
ICT	Information and Communications Technology
ICU	Intensive Care Unit
ID	Identity
IDE	Integrated Development Environments
IDK	Interface Development Kit Archive
IDP	Identity Provider
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IFAA	International Federation of Associations of Anatomists
IHE	Integrating the Healthcare Enterprise
IHI	Innovative Health Initiative
IMDRF	International Medical Device Regulators Forum
IMS	Infrastructure Management Services
INCF	International Neuroinformatics Coordinating Facility
INP	Input file
INRIA	Institut National de Recherche en Informatique et Automatique
IoT	Internet of Things
IPFS	InterPlanetary File System
IPR	Intellectual Property Rights
IPS	International Patient Summary
IQWiG	Institute for Quality and Efficiency in Healthcare
ISFET	Ion-sensitive field-effect transistor
ISO	International Standards Organisation
IST	<i>In silico</i> Trials
ISW_CoP	<i>In silico</i> World Community of Practice
IT	Information Technology
ITU	International Telecommunication Union
IVDR	<i>In vitro</i> diagnostics medical devices regulation
JIC	Joint Initiative Council for Global Health Informatics Standardization
JSON	JavaScript Object Notation
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KG	Knowledge Graph
KGML	KEGG Markup Language
KiSAO	Kinetic Simulation Algorithm Ontology
LAT	Local Activation Time
LLM	Large Language Model
LYN	Lynkeus SRL
M	Month
mCode	minimal Common Oncology Data Elements
MDCG	Medical Device Coordination Group

MDDT	Medical Device Development Tool
MDR	Medical Device Regulation
MDSW	Medical Device Software
ME/CFS	Myalgic encephalomyelitis/Chronic Fatigue Syndrome
MEE	Model Execution Environment
MIASE	Minimum Information About a Simulation Experiment
MIMIP	Model Inference and Parametrization
MINIMAR	Minimum Information for Medical AI Reporting
MIRIAM	Minimum Information Requested in the Annotation of biochemical Models
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MMSL	Multiscale Modeling and Simulation Language
MNC	Multi-national Companies
MOT	Model Object Type
MPI	Message Passing Interfaces
MRI	Magnetic Resonance Imaging
MS	Member State
MSF	Minimum Side-Fall Strength
MTB	Molecular Tumor Boards
MultiCellDS	MultiCellular Data Standard
MultiCellML	Multi-Cellular Modelling language
MVP	Minimal Viable Product
NCA	National Competent Authorities
NCAPR	National Competent Authorities for Pricing and Reimbursement
NCBI	National Center for Biotechnology Information
nD	n-dimensional
NeuroML	Neuroscience eXtensible Markup Language
NFS	Network File System
NGO	Non-governmental Organisation
NGS	Next generation sequencing
NICE	National Institute for Health and Care Excellence
NIFTI	Neuroimaging Informatics Technology Initiative
NIH	National Institutes of Health
NIST	US National Institute of Standards and Technology
NLP	Natural Language Processing
NEF	Neural Network Exchange Format
NoSQL	‘Not only’ Structured Query Language database
NPY	NumPY file format
NRRD	Nearly Raw Raster Data
NuML	Numerical Markup Language
OA	Open access
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OBI	Ontology for Biomedical Investigation
OBI	Ontology for Biomedical Investigation
ODE	Ordinary Differential Equation
OECD	Organization for Economic Cooperation and Development
OHDSI	Observational Health Data Sciences and Informatics
OIDC	Open ID Connect
OM	Operations and Monitoring
OMEX	Open Modelling Exchange
OMOP	Observational Medical Outcomes Partnership
ONNX	Open Neural Network Exchange
OpenAIRE	Open Access Infrastructure for Research in Europe
OpenCARP	Open Cardiac Arrhythmias Research Package
openEHR	open Electronic Health Records
OpenMINDS	Open Metadata Initiative for Neuroscience Data Structures
OpenVT	Open Virtual Tissues
ORCID	Open Researcher and Contributor Identifier

OSS	Open Source Software
PaaS	Platform as a Service
PACS	Picture Archiving And Communication System
PBPK	Physiology-based pharmacokinetics
PC	Project Coordinator
PCCP	Predetermined Change Control Plan
PCP	Pre-commercial procurement of innovation
PEA	Privacy Enforcement Authorities
PerMedCoE	Personalised Medicine Centre of Excellence
PET	Privacy-enhancing technology
PFA	Portable Format for Analytics
PFB	Portable Format for Bioinformatics
PGx	Pharmacogenomics
PharmML	Pharmacometrics Markup Language
PHD	personal health devices
PI SCHOOL	PI SCHOOL
PICAR	Population, Intervention, Comparison, Attributes of eligible Clinical Practice Guidelines, Recommendation
PICOR	Population, Intervention, Comparator. Outcome and Recommendation
PK	Public pseudonymization Key
PM	Project Manager
PMML	Predictive Model Markup Language
PoC	Proof of Concept
PoW	Proof of Work
PPI	Public Procurement of Innovation
PPP	Public-Private Partnership
PPRL	Privacy Preserving Record Linkage
PRE	Proxy Re-Encryption
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
Q1	Quarter 1
QALY	Quality Adjusted Life Years
QMUL	Queen Mary University of London
QnA	Question and Answer
QoI	Quantity of Interest
R&D	Research and Development
RAM	Random Access Memory
RCT	Randomized Control Trial
RDA	Research Data Alliance
RDF	Resource Description Framework
RDM	Research Data Management
RDP	Remote Desktop
RI	Research Infrastructure
RNG	Random Number Generator
RO	Research Object
ROI	Region of Interest
RoI	Return on Investment
RRO	Runtime Requirements Ontology
RWE	Real World Evidence
RWTH	Rhine-Westphalia Technical University of Aachen
SaaS	Software as a Service
SAIL	Secure Anonymised Information Linkage
SAM	Serviceable Available Market
SaMD	Software as a Medical Device
SBGN	Systems Biology Graphical Notation
SBML	Systems biology markup language
SBO	Systems Biology Ontology
SBOL-Visual	Systems Biology Open Language
SBRML	Systems Biology Results Markup Language

SDO	Standards Developing Organisation
SED-ML	Simulation experiment description markup language
SEIM	Security Information and Event Management
SEP	Standard Essential Patents
SEPIO	Scientific Evidence and Provenance Information Ontology
SePR	Secure eResearch Platform
SK	Secret (private) pseudonymization Key
SME	Small and Medium-sized Enterprises
SMPC	Secure Multi-Party Computation
SNOMED	Systematized Nomenclature of Medicine
SO	Standardized Output
SOM	Serviceable Obtainable Market
SOP	Standard Operating Procedure
SPDX	Software Package Data eXchange
SPE	Secure Processing Environment
SPS	SaMD Pre-Specifications
SRB	Single Resolution Board
SSH	Secure Shell Protocol
SSO	Single Sign-on
STAR*	Strategies to Reduce Transmission of Antimicrobial Resistant Bacteria
STEEPLE	Societal, Technical, Economic, Environmental, Political, Legal, and Ethical
STL	Stereolithography File Format
Stl	Stereolithography file format
STT	Specific Tech Tools Businesses
SV	Shapley Value
SWOT	Strengths, Weaknesses, Opportunities, Threats
TAM	Total Addressable Market
TC	Technical Committee
TCGA	Cancer Genome Atlas
TCL	Tool Command Language
TDM	Text and Data Mining
Teddy	Terminology for the Description of Dynamics
TEE	Trusted Execution Environment
TEHDAS	Towards European Health Data Space
TES	Task Execution Service
TIFF	Tagged Image File Format
TIR	Time In Range
TPLC	Total Product Lifecycle
TRL	Technology Readiness Level
TS	Technical Specification
TS4NFDI	Terminology Services 4 NFDI
TSD	Directive on the Protection of Trade Secrets
TTO	Technology Transfer Office
TTP	Trusted Third Party
TTSO	Tech Transfer Or Spinoff Businesses
UDI	Unique Device Identification
UDI-DI	Unique Device Identification – Device Identifier
UDI-PI	Unique Device Identification – Production Identifier
UI	User Interface
ULIÈGE	University of Liège
UNIBO	Alma Mater Studiorum – Università di Bologna
UP	Unitary Patent
UPC	Unified Patent Court
UQ	Uncertainty quantification
US	United States of America
USD	United States Dollars
USP	Unique selling proposition
UvA	Universiteit van Amsterdam

UVA	University of Virginia
VAEs	Variational Auto-Encoders
VDI	Virtual Desktop Infrastructure
VEP	Variant Effect Predictor
VPHi	Virtual Physiological Human Institute
VHT	Virtual Human Twin
VITO	Flemish Institute for Technological Research
VM	Virtual Machine
VR	Virtual reality
VTk	Visualization Toolkit
VV-40	Verification and Validation 40
VVUQ	Verification, Validation and Uncertainty Quantification
W3C	World Wide Web Consortium
WDL	Workflow Descriptive Language
WES	Workflow Execution Service
WHO	World health organisation
WIPO	World Intellectual Property Organization
WoT	Web of Things
WP	Work Package
WP29	Article 29 Working Party
XML	Extensible Markup Language
y.o.	Years old
Y1	Year 1
ZKP	Zero-Knowledge Proof

1 Genesis and validation of the roadmap

1.1 The EDITH-CSA project

The Virtual Human Twin (VHT) is an integrated multi-level, multi-time, and multi-discipline representation of quantitative human physiology and pathology. Its realisation through collaborative distributed knowledge and resource platforms is specifically designed to accelerate the development, integration, and adoption of patient-specific predictive computer models, which will be used as clinical decision support systems for personal health forecasting or as methodologies for the development and de-risking of personalised medical products. The **vision of EDITH** is to facilitate the realisation of the opportunities presented by VHTs for the benefit of patients, healthcare providers, regulatory bodies and industry, and other ecosystem actors, both within Europe and globally.

EDITH was a Coordination and Support Action (CSA) funded by the European Commission (EC), capitalising on the developments of digital technologies, employment of high-performance computing (HPC), availability and access to research and healthcare data in Europe, with the **mission** of creating a roadmap to go from the currently available resources (which often focus on a single organ or a single function system) to a data-driven and knowledge-driven fully integrated multi-scale and multi-organ whole-body VHT. EDITH-CSA facilitated this process by building an inclusive ecosystem driven by a consensus among the relevant European communities following RRI principles, and building a proof-of-concept infrastructure consisting of a data/model repository (and catalogue), as well as a simulation platform.

The objectives of the EDITH-CSA project were the following:

- To frame an ecosystem of Digital Twins in healthcare within the European Union (EU). EDITH is conducting a **mapping** of actors, initiatives, resources, and barriers in the Digital Twins, with the aim of ensuring adequate clinical representation and fostering the integration of all relevant stakeholders such as developers, technology and infrastructure providers, end-users, regulatory agencies, and Health Technology Assessment (HTA) bodies;
- To build a **roadmap** towards an integrated Virtual Human Twin, identifying the main research challenges and infrastructure needs, formulating clear policy recommendations, addressing interoperability, computability and health information integration, identifying implementation needs/barriers and developing a strategy for the clinical deployment of the VHT models and their uptake in personalised clinical decision-making.
- To develop a **federated and cloud-based repository** of Digital Twins in healthcare (data, models, workflows, algorithms, and good practices), pooling together existing resources across Europe and providing access to relevant existing data and model repositories. The ecosystem was leveraged to create a repository catalogue with available resources and recruit resources from the consortium and beyond. The ultimate goal for this repository was to understand the diversity of needs and expectations of the ecosystem in order to adequately inform the roadmap.
- To outline a **simulation platform** supporting the transition towards an integrated VHT implemented as a public infrastructure, providing a one-stop shop to design, develop, test, and validate single-organ Digital Twins and combine them with others for the integrated VHT models. Five use cases involving a range of diseases, organ systems and medical fields (cancer, skeletal diseases, cardiovascular and brain applications and applications in intensive care medicine) had been pre-selected to be developed as prototypes to show the added value of a simulation platform. In the course of the project, the decision was taken to build a proof-of-concept platform (though not required by the grant agreement) in order to test several choices and functionalities included in the roadmap recommendations.

1.2 Genesis of the vision and roadmap outline

The original vision as proposed in Viceconti et al. (2023)² was the result of work initiated in the grant preparation phase and continued through the review phase and the project execution. An articulated vision was expected by month 6 of the project (31/3/2023) and a first version of the entire roadmap³ by month 10 (31/7/2023). Given that challenging timeline, the work on the vision was prepared by the scientific community of experts with a proven track record in one of the various scientific disciplines and fields relevant to the VHT, in active collaboration with specific external experts including industrial

² <https://doi.org/10.1109/jbhi.2023.3323688>

³ EDITH CSA Deliverable 3.2: first draft of the VHT roadmap. Zenodo. <https://doi.org/10.5281/zenodo.8200955>

and clinical translation. Several public meetings and public draft consultations were organised in order to further validate the proposed concepts. The submission of the first draft of the VHT roadmap marked the end of the first phase of EDITH-CSA.

1.3 Ecosystem engagement activities

The list below captures the meetings organised by EDITH-CSA. The main discussion meetings were recurring online meetings with the consortium, the advisory board/group of stakeholder and public discussion meetings with (parts of) the ecosystem:

- EDITH consortium meetings
 - October 11, 2022, online: kick-off meeting;
 - October 2022-December 2024: Mapping, Vision, Repository/Platform and Sustainability WG meetings, each meeting on a biweekly basis;
 - November 29-30, 2022, Leuven;
 - January 30-31, 2023, Leuven;
 - April 16-18, 2024, Athens.
- EDITH advisory board meetings
 - December 2022-July 2023: Industry Advisory Board meeting every 2 weeks;
 - September 2023-November 2024: 5 Industry Advisory Board deep dives;
 - May 13 (intro), June 25, 26, 28 (deep dives), November 15 (recommendations), 2024: Advisory Group of Stakeholders meetings.
- Public meetings
 - May 16-17, 2023: EDITH Deep Thinkers Meeting Rome (100 experts covering all stakeholder groups);
 - March 24, May 2 & June 1, 2023: EDITH public discussion meetings (online);
 - December 22, 2023: VHT launch (organised by European Commission, onsite & online presence);
 - January 16-17, 2024: EDITH ecosystem meeting Paris;
 - April 17, 2024: EDITH public meeting with external contributors of use cases (online);
 - July 15-16, 2024: EDITH ecosystem meeting Amsterdam;
 - November 27, 29, 2024: EDITH public discussion meetings (online);

The public ecosystem meetings consisted of a combination of plenary sessions, strategy panel discussions and breakout sessions. The topics of the panel discussions and breakout sessions reflected the different aspects covered in this roadmap. Reports of the public meetings of Rome^{4,5}, Paris^{6,7} and Amsterdam^{8,9} are available on the EDITH-CSA website, along with the slide decks.

In addition to these meetings, over the entire duration of the CSA, EDITH/VHT-focused meetings with various communities of the ecosystem were held. These included meetings with clinical societies, scientific societies, research infrastructures, industrial partners, policy makers, regulatory actors and patient organisations. Furthermore, EDITH consortium members were invited to discuss the VHT initiative at numerous conferences, workshops and summer/winter schools in the fields of digital twins, *in silico* medicine, systems biology, wearables, HPC, AI, organ-on-chip, clinical science, ethics, regulatory science, and policy making. Additionally, a wide range of public events and community of practice meetings (*e.g.*, Avicenna Alliance, VPHi) have taken place where EDITH consortium members have presented the current status of the developing VHT vision and roadmap. Finally, key events, developments and opportunities were communicated in a monthly newsletter.

1.4 Writing of the roadmap

The written representation of the aforementioned meetings and discussions resulted in the current document. A first public document detailing the **vision and roadmap outline** was published in Zenodo and opened for comments and discussion¹⁰. This was followed by a **preliminary version of the first**

⁴ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH-DTM-Rome_Minutes.pdf

⁵ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH_DTM_Rome_annexes.zip

⁶ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH-EM-Paris_Minutes.pdf

⁷ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH_EM_Paris_annexes.zip

⁸ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH-EM-Amsterdam_Minutes.pdf

⁹ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH_EM_Amsterdam_annexes.zip

¹⁰ EDITH CSA Deliverable 3.1: Vision for the Virtual Human Twin and Roadmap Outline (2023). <https://doi.org/10.5281/zenodo.7796845>

draft of the roadmap, a preliminary version of the final draft of the roadmap and the draft version of the extended summary of the final roadmap.

- April-May 2023: vision and roadmap outline (via Zenodo and website form);
- June-July 2023: initial draft of the roadmap (via Zenodo and website form, or editable Google Docs);
- July-October 2024: comments on the first draft of the roadmap, leading to the final draft (via editable Google Docs);
- November-December 2024: extended summary of the roadmap (via editable Google Docs).

All inputs gathered from the activities mentioned above have been thoroughly discussed and processed in the roadmap. Hence, when the word ‘we’ is used in this roadmap, it refers to the entire community.

1.5 Validation of the roadmap

1.5.1 General validation approach: key principles

There is no well-defined process for the validation process for roadmaps, however, several elements were identified as **key principles** to the validation of roadmap building:

- Need identification: the roadmap’s ability to identify and address existing gaps in programs;
- Understanding: the ease of roadmap and roadmap content comprehension;
- Acceptability: the perception, satisfactoriness and organizational fit of the roadmap;
- Integration: the roadmap’s ability to fit into the current programming;
- Translation: measures the facilitators of translation such as appropriateness, context nuances and budget feasibility;
- Implementation: measures the possible execution of all sections and actions listed in the roadmap;
- Practicality: the ability of existing financial and human resources to implement the roadmap;
- Political buy-in: the political support for the implementation of RM (from individual organisations up to EU level).

The activities of the EDITH-CSA consortium and the approach taken to building and writing the VHT roadmap have been consistently based on these key principles.

1.5.2 Validation by an advisory group of stakeholders

During the second year of the project, the Advisory Group of Stakeholders (AGS) was activated. The AGS, through its composition of high-level and internationally renowned experts, is an important instrument for strengthening the VHT ecosystem, delivering **high-level feedback and validation**, and, in doing so, **generating political buy-in for the VHT** roadmap and the proposed activities. The AGS was composed of high-level experts from academia, clinics, industry (large companies and start-ups), trade associations, regulatory agencies and notified bodies. The expertise of the AGS members covered the technological (academic & industrial), ethical, legal and social (ELSI) aspects related to VHT. The members of the AGS were the following:

- Valentina Strammiello, European Patient’s Forum
- Mariano Vasquez, ELEM Biotech
- Ger Janssen, Philips
- Annabel Seebohm, COCIR
- Peter Aulbach, Siemens Healthineers (represented by Lance Ladic)
- Julia Neguer, Dassault Systèmes (replaced by Jean Colombel)
- Patrick Boisseau, MedTechEurope
- Robert Madjno, TüvSüd (represented by Briane Laruy and Surash Surash)
- Gernot Marx, Uniklinik Aachen
- Fidelia Cascini, Catholic University of the Sacred Heart, Rome
- Michael Benson, Karolinska Institutet
- Julio Saez-Rodriguez, University of Heidelberg
- Paola Grosso, University of Amsterdam
- Maxime Sermesant, INRIA
- Signe Mezinska, University of Latvia
- Maria Benedetti, University of Bologna

The first meeting of the AGS took place online on May 13th 2024, starting with a general introduction to the EDITH project, its consortium and main goals. This was followed by a discussion of the VHT roadmap approach, structure & subject matter. This first meeting was followed by three deep dive meetings into the VHT science, technology & infrastructure (25/6/2024), the Ethical, Legal and Social issues for VHT (26/6/2024) and the uptake, users & sustainability (28/6/2024). The last meeting of the AGS in the scope of the EDITH project (15/11/2024) was dedicated to the roadmap and its recommendations. The basis of the discussion was the extended summary of the roadmap, a high-level document aimed at policy makers explaining the rationale, reality and realization of the VHT initiative and infrastructure. All AGS meeting minutes are available on the EDITH website¹¹. Below we briefly discuss each meeting by providing the agenda and the summary of the recommendations that were made.

1.5.3 From an innovation ecosystem to an innovation framework

The essential starting point for ensuring validation is the connection, interrogation and collaboration with an active and engaged ecosystem, as well as a thorough view on the state of the art of the relevant VHT actors and building blocks. Through its mapping, stakeholder engagement, and innovative knowledge sourcing tool, EDITH-CSA addressed gaps in visibility, fostered collaboration, and laid a strong foundation for the sustainable development and adoption of the European VHT.

The **innovation ecosystem** identified in this deliverable is the basis upon which the validated VHT roadmap will define the VHT's innovation framework:

- By having **identified the stakeholders and brought together the ecosystem**, EDITH-CSA has accomplished the first step of the successful realisation of an inclusive and fit-for-purpose VHT. Having all stakeholders on board, sharing a vision and contributing building blocks is the best guarantee for a swift uptake and committed user community.
- By having **identified the needs and expectations** of each of the stakeholders, EDITH-CSA has provided the future developers of the VHT infrastructure with clear guidance on how to realize an inclusive and sustainable VHT advanced simulation platform and related services.
- By having **actively engaged with the entire ecosystem** in a variety of activities throughout the roadmapping process, the EDITH-CSA consortium has addressed all the elements that constitute the validation of the proposed roadmap. For the future development of the VHT, it will be absolutely essential to continue to invest in the ecosystem development and engagement – expanding it from the global level to the national and regional level within the different EU member states.

In conclusion, the VHT ecosystem is an essential prerequisite for executing the vision detailed in this roadmap and building a sustainable and active **innovation framework** that is able to deliver on the promise of the Virtual Human Twin.

¹¹ <https://www.edith-csa.eu/advisory-boards/>

2 Structure of the Roadmap

2.1 From need statement to sustainability

The roadmap is structured along the following 5 parts

- Need and vision
- Technology
- Infrastructure
- Standards, regulatory, ethical, legal and social issues
- Users, uptake and sustainability

This structure allowed us to start from the ecosystem's needs and expectations before defining a vision and identifying the necessary building blocks to realise that vision. When defining the building blocks, the state of the art is always considered, not only to avoid reinventing the wheel, but also – and more importantly – to be able to interoperate with the wide range of already available public and private building blocks and initiatives. At the end of the roadmap, in part 5, we return to the stakeholders and discuss sustainability. The final part, part 6, provides a comprehensive list of the recommendations from each of the 5 parts, a tentative timeline for their development, and a series of tangible actions for the different stakeholders in the VHT ecosystem.

2.2 VHT roadmap for stakeholders

Given our choice to make this roadmap a comprehensive document that also includes many technical details about technologies, standards and EU regulations, we have decided to produce 2 more versions of this roadmap. An extended summary of this roadmap will be published alongside this comprehensive roadmap. This **extended summary** is intended as a **policy brief**, created for policy makers and anyone interested in understanding the need for and the multi-faceted nature of the VHT, as well as the tangible steps (recommendations) that need to be taken to realize the VHT. The extended summary will be translated into French, Spanish, Italian, German, Dutch, Swedish and Polish languages.

A final version of this roadmap will appear towards the end of 2025 in the form of a **popular science book** on digital twins in healthcare. This, too, will cover all the parts of this roadmap but in a light and accessible tone. The aim of the book is to provide information on digital twins, their technological, standards, and regulatory needs, along with ELSI considerations.

2.3 Success stories, use cases and user stories

To illustrate the various sections and concepts, a variety of examples have been included in this roadmap. There are four categories of examples:

- Success stories: these are examples of currently existing Digital Twins, communities, initiatives, infrastructures etc. These are short descriptions, highlighting a particular aspect of the example to illustrate a specific section in the roadmap.
- EDITH developments: these are examples of tools or demonstrators developed during the EDITH-CSA project with the aim of informing the roadmap.
- Use cases: these are examples of currently existing Digital Twin solutions that have gone through all phases of research and development and that are currently present in (or being transitioned into) clinical practice. These use cases will come back in every part of the roadmap (technology, infrastructure, ELSI, uptake), providing a complete picture of a VHT application.
- User stories: these are fictitious stories that shed light on a possible future scenario for VHT development, from the angle of different stakeholders. These user stories are inspired by actual applications and ongoing developments, making the realisation of said stories likely in a future where the VHT roadmap has been implemented.

All these examples are included throughout the roadmap and are recognisable by the coloured box they are presented in. The header of the box clearly states what type of example it is (success, story, use case or user story). For success stories and use cases, references are provided for further reading.

Box 1: Success story - example

Success story: Name

Here you can read the summary of the example.

PART 1:

**FROM DIGITAL TWINS IN HEALTHCARE TO THE
VIRTUAL HUMAN TWIN**

3 Definitions

3.1 Digital Twins in Healthcare

A **Digital Twin (DT)** is the virtual representation of a physical object or system across its lifecycle. It uses real-time data and other sources to enable learning, reasoning, and dynamically recalibrating for monitoring, diagnostics, and prognostics. When applied in the context of health and care, the elements of real-time read-out and dynamic recalibration are relevant for specific applications and contexts (such as ICU patient care and deep brain stimulation), but for many other applications these elements are not feasible and/or needed. Hence, for most applications in health and care, the interactions between the Digital Twin and its physical counterpart are realized through a human-in-the-loop¹². Hence, the concept of **Digital Twins in healthcare**, is extended to the direct use of individual-specific models for the prevention, prediction, screening, diagnosis and treatment of a disease, as well as the evaluation, optimization, selection and personalisation of intervention options.

Digital Twins can focus on pathophysiological processes at **multiple time and length scales**, from a single cell over tissues to the organ or system level, or **treatment strategies** (devices, drugs or advanced therapies). In addition, Digital Twins can also cover **manufacturing processes** related to the aforementioned therapeutic strategies, or even entire **medical facilities**. Henceforth, when the term DT is used, it is understood to be one of the aforementioned applications.

The key potential in health and care of this technology is related to targeted prevention, tailored clinical pathways, and supporting healthcare professionals in virtual environments. Examples include the implementation of clinical trials for medicines and devices, medical training, surgical intervention planning, and several other potential use cases in virtual world environments.

3.2 Digital Twins in Healthcare – core components

The Digital Twin system consists of three core components: hardware, data management middleware, and software. Together, these elements form the backbone of the DT system, each playing a vital and interconnected role. Each of these components—hardware, middleware, and software—contributes uniquely to the Digital Twin ecosystem in healthcare, working in harmony to bridge the physical and digital worlds and unlocking opportunities for personalized medicine and improved patient care.

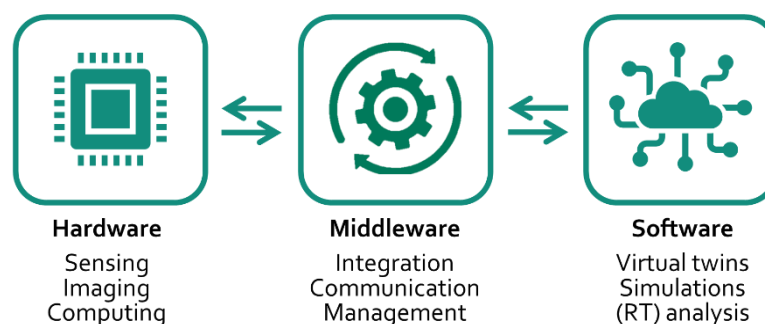


Figure 1: key components of Digital Twins

3.2.1 Hardware Components

Hardware serves as the physical layer of a Digital Twin system, comprising medical imaging, sensors, actuators, devices, and computational infrastructure that capture and process real-world data. These components provide the critical foundation for creating digital replicas.

For example, consider a patient undergoing treatment for a cardiovascular condition. Wearable sensors, such as electrocardiogram (ECG) monitors, blood pressure cuffs, and oxygen saturation trackers, continuously collect data about the patient's physiological parameters whereas medical images provide insights into the overall status of the system. These hardware components act as the inputs of the Digital Twin, providing data that reflects the patient's cardiovascular function and overall health status.

¹² <https://doi.org/10.17226/26894>

3.2.2 Middleware Components

Middleware acts as an intermediary between hardware and software, enabling seamless data communication, integration, and management. This layer ensures that data collected from the hardware is processed, standardized, and made accessible to the software for further analysis and simulation.

In a hospital setting, for instance, multiple patients in the intensive care unit (ICU) are monitored using various medical devices that measure parameters such as heart rate, respiration, and glucose levels. Middleware aggregates and organizes this data, ensuring that it is standardized and consistent for use by the Digital Twin's software layer. This ensures that clinicians can access accurate and integrated data to support treatment decisions.

3.2.3 Software Components

Software represents the analytical and predictive layer of the Digital Twin system. It leverages pre-existing knowledge on (patho)physiology along with data provided by the hardware and middleware to create virtual models, conduct simulations, and perform predictive analyses to support informed decision-making.

Take the example of personalized treatment planning for a cancer patient. Digital Twin software can create a virtual model of the patient's tumour using imaging data, biopsy results, genetic screening information and sensor-collected biomarkers. By simulating how the tumour responds to different therapies, such as chemotherapy or radiation, the software helps oncologists identify the most effective treatment plan. This allows clinicians to optimize patient outcomes while minimizing adverse effects.

3.3 Digital Twin models – key concepts

The management of human health in its broadest sense requires decision-makers to make **well-informed decisions that may affect the health status of single or groups of human beings** (hereinafter generically called *reference population*). Examples of this include clinicians making decisions on personalised therapeutic strategies for a patient; researchers making decisions on possible druggable targets to pursue in basic biomedical research; healthcare authority managers planning specific policies and optimising existing resources; biomedical companies seeking to refine, reduce and partially replace animal and human experimentation for the regulatory approval of new products; *etc.* This decision-making process usually involves the quantification of specific constructs that represent such health status called outcomes, with selected metrics called **Quantity of Interest (QoI)**, and then observing how such QoI develops in time, due to variations of internal (*e.g.*, body weight) or external (*e.g.*, exposure to pollutants) conditions, or because of intentional interventions. The **Context of Use (CoU)** defines how the QoI informs a specific decision-making process relevant to human health and under which specific conditions such a process occurs. Typical VHT-related CoUs¹³ include diagnosis^{14,15,16}, prognosis^{17,18}, personal health forecasting, patient stratification, clinical decision support¹⁹, development of medical products and *in silico* clinical trials.

QoIs are usually measured, either directly on human volunteers or patients during clinical trials, or indirectly on surrogates such as animal or *in vitro* experimental models. However, these studies pose a long list of practical, ethical, legal and socioeconomic challenges and contribute to healthcare services' high costs and limited capacity. Thus, there is intense ongoing research on developing new technologies that can refine, reduce and partially replace the need for experimental measurements to estimate the QoIs necessary to support decision-making within specific CoUs.

The software component of a Digital Twin is a computer simulation that predicts (as opposed to measuring experimentally). The QoI necessary to support decision-making within a specific CoU in healthcare Here, “computer simulation” refers to any software (including AI-based software) capable of predicting specific outputs given certain inputs. DTs can be predominantly knowledge-driven

¹³ Vicenconti *et al.* IEEE J Biomed Health Inform. 2021; 25(10):3977-3982. <https://pubmed.ncbi.nlm.nih.gov/34161248>

¹⁴ Melis *et al.* J. Biomech. 2019; 90, 24-32. <https://doi.org/10.1016/j.jbiomech.2019.04.019>

¹⁵ Gosling *et al.* European Heart Journal - Digital Health, 2022;3(3),481–488. <https://doi.org/10.1093/ehjdh/ztac045>

¹⁶ Lungu *et al.* Pulmonary Circulation 2016;6(2),181-190. <https://doi.org/10.1086/686020>

¹⁷ <https://www.ansys.com/advantage-magazine/volume-xv-issue-3-2021/simulation-and-high-performance-computing-reduce-fracture-risk-in-osteoporotic-patients>

¹⁸ Teixeira *et al.* Biomech Model Mechanobiol 2020;19, 2413–2431. <https://doi.org/10.1007/s10237-020-01351-2>

¹⁹ Lin *et al.* Computer Methods and Programs in Biomedicine 2011; 102(2),192-205. <https://doi.org/10.1016/j.cmpb.2010.12.008>

predictive models built using existing – validated or hypothesised – knowledge about physics, chemistry and (patho)physiology. Alternatively, they can be predominantly data-driven models built from large volumes of data using statistical modelling or artificial intelligence (AI) / machine learning (ML) techniques, or any combination²⁰. It should be stressed that while the DTs can predict QoIs that are difficult or impossible to measure, they can do a lot more in general: they can predict how QoIs will evolve in time, how they will change depending on external actions, *etc.*

One of the most important features of Digital Twins is the accuracy with which they predict such quantities; thus, DTs can be divided into three broad categories:

- **Generic DT**, for which the expected accuracy is that the predicted value is within the range of the values measured experimentally in the reference population;
- **Population-specific DT**, for which the expected accuracy is that the predicted value is sufficiently close to some central property (typically mean or median) of the range of the values measured experimentally in the reference population;
- **Subject-specific DT**, for which the expected accuracy is that the predicted value is sufficiently close to the value measured experimentally in each individual in the reference population.

By “sufficiently close”, we mean that the predictive accuracy of the DT is sufficient for its purpose as defined in the CoU. Thus, the same DT can be sufficiently accurate for one CoU and insufficiently accurate for another. Therefore, it is fundamental to define the target accuracy of the DT when describing its CoU.

Currently, most DTs are designed to predict just one or a small number of QoIs with the necessary accuracy only in a narrowly defined reference population (*e.g.*, women over 55 with osteoporosis and no other conditions). This is because to develop credible DTs, we need large volumes of detailed empirical observations and/or reliable mechanistic knowledge of the physiology and pathology/pathophysiology of the organs, tissues, and cells involved, as well as the mechanism of action of any intervention involved. Because of gaps in knowledge and data, the only way to manage this complexity today is to narrow the scope of the DT, focusing on a particular process affecting a minimal portion of the human body, and to be used for narrowly defined CoUs. While these narrowly focused DTs are extremely useful in specific cases, the time and cost required to develop DTs with a broader scope and wider applicability are currently prohibitive due to challenges in addressing gaps in our knowledge and linking and reusing existing developments. The Virtual Human Twin (VHT) can provide such a framework, actively identifying and filling gaps in our knowledge, as well as developing solutions for CoUs that can be provided with existing knowledge. Hence, the VHT can enable a new generation of DTs, capable of predicting any QoI necessary for any relevant CoU and reference population.

3.4 The life cycle of Digital Twin models

The development of a DT model can be a long and cumbersome yet highly rewarding process. It starts with the **identification of the clinical needs**, expressed through epidemiological evidence or any other methodology that quantifies the limits of the current standard of care. These needs may refer to healthcare challenges for which there is no existing clinical solution, but could also refer to aspects of health and care that provide significant improvement in patient care or outcomes, or a significant cost reduction. HeartFlow is one of the first DTs adopted in clinical practice²¹, addressing a clear clinical need. Even though there is universal consensus among cardiologists that the best way to choose the most appropriate treatment for coronary stenosis is a Fractional Flow Reserve (FFR) measurement obtained through an invasive diagnostic test, only 20% of the UK patients with this condition are treated based on an FFR. HeartFlow provides a quantification of the FFR based on medical images. Another clinical context where a mature DT is available, is fracture risk prediction in osteoporotic patients, where the standard of care requires specialists to decide whether to treat a patient with anti-osteoporotic drugs using the current gold standard X-ray based bone mineral density measurement as a predictor of hip fracture risk. With this risk predictor, around one-third of the patients are not treated; of those, around 50% will experience a hip fracture in the following five years. Considering that current treatment

²⁰ <https://doi.org/10.5281/zenodo.8064147>

²¹ Rasoul H et al., Clin Med (Lond). 2021;21(2):90-95. <https://doi.org/10.7861/clinmed.2020-0691>

can reduce the incidence of hip fracture by around 50%, with a better risk predictor, up to 7.5% of all hip fractures (60,000 per year only in Europe) could be avoided. The Bologna Biomechanical CT-Hip (BBCT-Hip) DT estimates fracture risk based on computer modelling and simulation on personalised patient data considering a wide range of fall scenarios²².

In addition, DTs can play a significant role in **addressing scientific biomedical needs** through the acquisition of basic knowledge, *e.g.* in understanding the multiscale microstructure of tissue organisation in various organs. These DTs do not necessarily have a direct clinical applicability but enable others to apply the obtained knowledge.

The second step is **determining the causal relationship** between the QoI (hereinafter generically referred to as the DT's outputs) and the parameters that control it (hereinafter generically referred to as the DT's inputs). Living organisms are *entangled*, meaning that each internal state variable depends on several other internal state variables. Of those, some variables have a greater effect on the QoI than others. Ideally, for each QoI, a sensitivity analysis should be run to understand how the variation of any other possible QoIs in the human body affects it. Since such a systematic is impossible, we use the available causal knowledge about the human body's physics, chemistry and (patho)physiology to identify the minimum set of inputs that would allow a reasonably accurate prediction of the desired outputs. In some cases, the desired input cannot be measured on a patient-specific basis. In such cases, we may build a statistical model for such a quantity that describes how it varies across the population of interest, possibly as a function of the inputs that can be measured for each patient. If this is not possible *e.g.*, due to the lack of information, we may vary the desired input in the range of values observed in the reference population and study its effect on the QoIs.

When the available causal knowledge is insufficient to build a reliable predictor, data-driven modelling techniques can be used to identify the best possible predictor from all available inputs. **Training a predictor** requires a large collection of data, both in depth (data from many diverse patients are required to train the predictor) and breadth (as we do not know *a priori* which quantities govern the QoI, we need to explore as many as possible). Prospectively, the creation of a DT will in many cases require improving causal knowledge, for example by eliminating some competing mechanistic hypotheses or complementing existing hypotheses, in which pre-stages of the DT may assist by guiding data acquisition.

The third step is **model implementation**. This is essentially a software development process that must be performed with the highest possible quality assurance. A key factor here is the availability of accurate input data to build benchmark problems used in the verification of the solvers. Another aspect is the definition of the DT's execution environment. Depending on the nature of the data (patient-related health data), there might be ethical-legal constraints requiring the data to be stored only at specific locations and under certain levels of cybersecurity. Depending on the model implementation, there might be computational requirements that impose that the model executes only on specific computers or in specific countries. The selection of the software for running the DT (either developed in-house or using existing commercial licenses) is important in light of the collaborative nature of the VHT. Given the expected complexity of a DT, software licences should be chosen to promote collaborative development and allows its use in early development stages. Metadata is essential to clearly define which license has been used and which use restrictions apply.

The fourth step is the development of all necessary **pre-processing and post-processing tools**. Pre-processing tools are those that extract the necessary DT inputs from available data and format them for use in the DT. For instance, we might need the volume of a tumour, which can be measured on a 3D MRI dataset, but only once the tumour is segmented in the images. The accuracy and the degree of automation of pre-processing tools are critical as excellent DTs can be poorly informed by sub-optimal pre-processed data which strongly reduces their predictive capabilities. Post-processing tools are required when the DT's output is not directly the QoI required to optimally support the clinical decision-making or knowledge-finding process, and for their clear presentation and/or visualization. Optimal use

²² Keaveny TM *et al.*, Osteoporos Int. 2020;31(6):1025-1048. <https://doi.org/10.1007/s00198-020-05384-2>.

of a DT is strongly favoured by embedding it in a well calibrated pipeline of data acquisition modalities, pre-processing, DT and post-processing.

The fifth and most important step is **model credibility assessment**. This rigorous process, dubbed Verification, Validation and Uncertainty Quantification (VVUQ), requires data from tightly controlled and highly qualitative experiments to conduct the validation. VVUQ is combined with applicability analysis to assess the relevance of the validation evidence to support using the model for its CoU. Once the technical validation is completed, additional clinical validation might be required. This should be done independently from those who developed the DT through the employment of data from prospective clinical studies. However, in some cases, the regulator may accept studies on (retrospective) registry data (as far as publicly available) as evidence of clinical validity. In the regulatory space, there is also an ongoing discussion on the possibility of certifying a DT by allowing the regulators to conduct validation studies against (publicly) available experimental data. In some jurisdictions, public access to datasets is not available, however, there are controlled access datasets available *e.g.* for those whose planned study is approved by an ethics committee for the purposes of obtaining regulatory approval. On a more general note, datasets that are publicly available typically consider a single organ, and therefore anonymisation may be approved given that the risk of re-identification is low. However, as more complex DTs are developed, they will require linking information from multiple organs, electronic health records etc, which can substantially increase the risk of re-identification even if “anonymised”. So in the future, it is unclear if such datasets can be made publicly available. To facilitate the process of credibility assessment, the *In Silico* World community of practice together with the VPH Institute and the Avicenna Alliance have drafted a proposal for a Good Simulation Practice²³ (GSP), in analogy to the Good Clinical Practice, as the go-to document for building credible *in silico* models in healthcare.

The last step is the **provision of (clinical) access**. A DT can be made available to the end-users as software embedded in medical imaging consoles, installable software, software-as-a-service, software as a medical device, *etc.* This is also related to business models to make the DT widely available, which can include not-for-profit modalities (for example, clinical end-users funding the further development of specialised DTs by not-for-profit organisations).

The DT life cycle is a **continuous and iterative process** where new insights and data (from *e.g.*, research or clinical deployment) can be used to further refine DTs over time.

3.5 The Virtual Human Twin

DTs in Healthcare, also indicated with the term Virtual Human Twins (VHTs), hold great potential in advancing personalised care, a core priority for the European Union. They will accelerate tangible benefits for citizens and patients, while also sustaining and advancing EU science and technology. Virtual Human Twin initiatives exist across many Member States. Academia, large industry, and SME innovators have already developed different versions of them in collaboration with clinicians, healthcare providers, and other stakeholders. However, even if the number of interested stakeholders is growing, the ecosystem remains highly fragmented in the EU.

The **Virtual Human Twin (VHT) initiative aims to tackle this fragmentation** on all levels. Conceptually, the VHT is to be a systematic, ever-growing digital, and quantitative representation of the actionable knowledge available on human pathophysiology. This entails, amongst others, the creation of a **federated public infrastructure** that will enable the pooling and linking of resources and assets (data, models, algorithms, computing power, storage etc.), facilitating the development of Digital Twins in healthcare and the assessment of their credibility. This will require **continued support** for advanced research, technology development and platform operation, leveraging the power of supercomputers and Digital Twin technologies, including AI, and supporting interoperability, integration and the scaling up of DT-based solutions. These developments need to take place in full **compliance with EU values and rules** related to privacy, safety and security. In addition, realising the

²³ <http://doi.org/10.1007/978-3-031-48284-7>

VHT will require an **engaged ecosystem**, which fosters inclusion and collaboration among diverse stakeholders.

To ensure maximal clarity for the reader of this roadmap, henceforth, the document will use the concept of DT for individual twins, whereas the concept of VHT will be used to denote the global initiative, its ecosystem and infrastructure.

4 Global Trends in the European Context

4.1 Trend 1: Escalating Healthcare Costs, Workforce Shortage and the Need for Efficiency

Globally, healthcare costs are rising at an unsustainable pace, straining healthcare systems and economies. This surge in expenditures is driven by multiple factors, including:

- **Ageing populations:** as life expectancy increases, the proportion of older adults in the global population is growing²⁴, leading to a higher demand for healthcare services. Older individuals often have more complex healthcare needs, contributing to higher healthcare costs.
- **Chronic and complex diseases:** the prevalence of chronic and complex diseases, such as cancer, cardiovascular disease, and diabetes, is increasing worldwide²⁵. Managing these conditions often requires managing multiple morbidities, as well as long-term and costly care, therefore placing significant financial burdens on healthcare systems.
- **Advances in medical technology:** while advancements in medical technology have led to improved treatments and diagnostic tools, these innovations often come with high price tags, contributing to the overall rise in healthcare costs.
- **Workforce shortages:** many countries are facing a growing shortage of healthcare workers, exacerbated by an ageing workforce²⁶. This shortage places additional strain on existing personnel, reduces system efficiency, and increases the cost of delivering care due to over-reliance on temporary staff or outsourcing.

The escalating healthcare costs necessitate a shift towards more efficient healthcare systems. The VHT can strongly contribute to a potential solution by enabling:

- **Prevention and early disease detection:** through continuous monitoring and analysis of individual health data, the VHT could facilitate the early detection of disease risk biomarkers and enable prevention and early treatment, potentially reducing the need for costly treatments later on. For instance, the MyDigiTwin project²⁷ provides a platform where patients' self-reported data is combined with data from their healthcare providers, and compared through AI-based algorithms with existing records, allowing patients to check and monitor their health and identify possible risks of various cardiovascular diseases.
- **Personalized treatment optimization:** by creating personalized models of individual patients, the VHT could help clinicians optimize treatment plans, reducing the likelihood of adverse reactions and improving the effectiveness of therapies, thus leading to cost savings. A wide range of DT-based surgical planning tools are currently in clinical use where critical choices (e.g. type and placement of stents) can be tested *in silico*.
- **Optimising development and delivery processes:** by creating virtual populations, the VHT could facilitate *in silico* clinical trials to complement and optimise *in vivo* clinical trials. Models of healthcare workflows and resource distribution (e.g. bed allocation), allow to optimise the overall logistics of the healthcare process.

4.2 Trend 2: The Imperative for Personalized Medicine

There is a growing need for personalised medicine, a paradigm that tailors medical care to individual patients based on their unique characteristics. Healthcare industry is transitioning towards this paradigm, Personalized medicine aims to move away from a one-size-fits-all approach to healthcare which acknowledges that individuals respond differently to treatments and interventions. This shift is driven by the understanding of the following elements:

- **Genetic and structural variation:** individuals possess unique genetic variations as well as structural variations (particularly in the brain, such as the connectome) that influence their susceptibility to diseases and their responses to medications.
- **Lifestyle factors and social determinants:** lifestyle factors, such as diet, exercise and social determinants, such as income and education also play a significant role in health and disease.

²⁴ <https://www.oecd.org/en/topics/policy-issues/the-future-of-health-systems.html>

²⁵ OECD (2023), <https://doi.org/10.1787/1e53cf80-en>.

²⁶ OECD/European Commission (2024), Health at a Glance: Europe 2024: State of Health in the EU Cycle, OECD Publishing, Paris,

²⁷ <https://www.mydigitwin.nl/>

- **Exposome:** environmental exposures such as pollution and the effects of climate change have been shown to affect patients' health.
- **Individual health history:** each patient has a unique health history that informs their current and future healthcare needs.

The VHT can serve as a powerful tool for strongly accelerating innovations required for advancing personalized medicine by:

- **Integrating multi-modal data:** the VHT can facilitate the seamless and automatic integration of data from various sources, including electronic health records, genetic tests, wearable sensors, and imaging studies, to create a comprehensive and personalized model of each patient.
- **Simulating treatment responses:** Using the personalized model, clinicians can virtually test on a computer how a patient might respond to different treatment options, enabling them to select the most effective and personalized approach.
- **Monitoring individual responses:** The VHT can continuously monitor patient responses to treatments and interventions, allowing for adjustments to treatment plans based on individual needs.

4.3 Trend 3: The Rise of Digitalization in Healthcare

The healthcare system is experiencing rapid digitalization, with the adoption of electronic health records (EHRs), telehealth, mobile health (mHealth) applications, wearable sensors, and AI. This digital transformation offers numerous opportunities to improve healthcare delivery, including:

- **Enhanced data collection and analysis:** digital technologies, in combination with advanced computing infrastructures and services, facilitate the collection, storage, and analysis of vast amounts of health data, providing insights into population health trends and individual patient characteristics, leading to more informed decision-making.
- **Improved patient engagement:** mHealth apps and wearable sensors empower patients to take a more active role in managing their own health, promoting self-monitoring and adherence to treatment plans.
- **Improved access to care:** telehealth and remote patient monitoring technologies enable healthcare providers to monitor patients outside of traditional clinical settings, improving access to care and reducing healthcare costs.

The VHT can effectively leverage advancements in digital healthcare in a number of key ways:

- **Advancing health and ehealth research:** the VHT and its infrastructure will allow the pooling and linking of resources and assets (data, models, algorithms, computing power, storage etc.), facilitating the development of Digital Twins in healthcare and the assessment of their credibility.
- **Accelerated therapy development:** patient populations created using the VHT can simulate clinical trials in a virtual environment. This application could potentially accelerate drug and device development, and as a result, reduce costs associated with traditional clinical trials.
- **Holistic view of patient health:** by integrating data from various digital health sources, the VHT can create a comprehensive, and dynamic view of each patient's health status.
- **Predictive analytics and proactive Care:** AI algorithms can be incorporated into the VHT to analyse patterns in individual health data, predicting disease risks and treatment outcomes. This predictive capability allows for more proactive and personalized healthcare interventions.

4.4 Trend 4: The EU Digital Agenda and the European Life Sciences Research Infrastructure

The EU is actively promoting the development of digital health technologies through initiatives such as the EU Digital Agenda. These initiatives focus on:

- **Fostering innovation in digital health:** the EU supports research and development in digital health technologies, including initiatives related to the VHT.
- **Developing digital health infrastructures:** the EU is investing in creating digital health infrastructures, such as those established under the European Health Data Space (EHDS), which

will support the secure and ethical sharing of health data. This infrastructure will be essential for developing and deploying the VHT effectively.

- **Organizing digital knowledge:** the EU is developing infrastructures for knowledge organisations such as EBRAINS and EUCAIM, which integrate and interoperate domain-specific data, models, and/or tools, as well as providing the operational basis of VHT.
- **Establishing regulatory frameworks:** the EU has been developing regulatory frameworks, on top of the General Data Protection Regulation (GDPR), such as the AI Act, the Data Governance Act and the European Health Data Space (EHDS), also to ensure the responsible and ethical use of digital health technologies, including the VHT.

The European life sciences research infrastructures, such as the European infrastructure for translational medicine (EATRIS) and the European life sciences infrastructure (ELIXIR), can play a vital role in supporting VHT development. These infrastructures provide access to:

- **Data and biobanks:** researchers can access a wide range of health data and biological samples to develop and validate VHT models. For instance, ELIXIR provides access to biological data, analysis tools, and computational resources, supporting research in fields such as genomics, proteomics, and systems biology. The VHT should utilize ELIXIR's resources to access and analyse large datasets, develop sophisticated models of human biology, and accelerate the development of VHT applications.
- **Computational resources:** advanced computing infrastructures are essential for running complex VHT simulations. The EuroHPC, Joint Undertaking was established to coordinate efforts between the EU and the participating countries for pooling resources and leading European supercomputing initiatives.
- **Expertise and collaboration:** research infrastructures foster collaboration between scientists from different disciplines and between stakeholders, which is crucial for the interdisciplinary and intersectoral nature of VHT research. For instance, EATRIS supports the translation of research discoveries into clinical practice, facilitating collaboration between researchers, clinicians, and industry partners. The VHT leverage expertise and infrastructure from EATRIS and other life sciences research infrastructures to bridge the gap between research and clinical implementation, ensuring that VHT-based innovations reach patients and healthcare providers effectively.

4.5 Trend 5: Contributing to the Sustainable Development Goals

The United Nations Sustainable Development Goals (SDGs) outline a global agenda for achieving a more sustainable and equitable future²⁸.

The VHT can play a role in achieving several SDGs, particularly those related to:

- **Good health and well-being (SDG 3):** by improving patient care, overall experiences and healthcare outcomes, promoting preventative care, and enabling personalized medicine, the VHT can contribute to ensuring healthy lives and well-being for all ages.
- **Gender equality (SDG5):** the VHT can be used to study the effects of sex differences on health and disease, and to develop more effective treatments for all patients.. *In silico* clinical trials can be used to design or augment clinical trials ensuring an adequate representation of the entire population.
- **Reduced inequalities (SDG 10):** the VHT can contribute to ensuring equitable access to healthcare by providing personalized and cost-effective solutions that are accessible to individuals regardless of their socioeconomic status or geographic location. For instance, the VHT can provide a virtual training environment to novice clinicians or non-clinicians who provide care at remote locations that otherwise have limited or no access to state-of-the-art information and healthcare.

²⁸ <https://sdgs.un.org/goals>

5 Vision for the Virtual Human Twin

5.1 The vision for the VHT

Convergence of global healthcare needs and trends is exposing the limits of traditional approaches to innovation, which are often reactive and one-size-fits-all, and provides a strong rationale for the development of the VHT. The **VHT offers a transformative approach to healthcare**, addressing the challenges of rising healthcare costs, the demand for personalized medicine, and the need for more efficient and effective healthcare systems. By leveraging the power of digitalization, aligning with EU policy initiatives, the VHT holds the promise of a healthier and more sustainable future for individuals and societies worldwide.

At the heart of the VHT initiative lies the **transformative power of *in silico* technologies**, including computer modelling and simulation as well as artificial intelligence. These are supported by the **indispensable hardware developments** related to computational infrastructure as well as data generation, storage and transport technologies. By creating virtual representations of human physiology and pathology through the use of human-relevant and patient-specific data, researchers and clinicians can gain unprecedented scientific insights into the complexities of the human body and disease progression across length scales, time scales, and organ systems. These insights can then be leveraged towards providing personalized, predictive and preventive care, accelerating medical research and innovation, and improving support for clinicians, healthcare providers and patients. It is important to remark that both the gathering of new knowledge as well as its application in improving clinical care, are important objectives of the VHT.

The ambitious vision of the VHT requires a collaborative effort that transcends national boundaries, including the **establishment of a European VHT infrastructure** in order to achieve its key goals, including:

- **Pooling Resources and Expertise:** Europe possesses a wealth of scientific, clinical and industrial expertise, R&D and healthcare organisations, and healthcare data. A European infrastructure can effectively pool these resources, facilitating collaboration and knowledge sharing among diverse stakeholders, including researchers, technology developers, clinicians, patients, regulators and policymakers.
- **Data Sharing and Interoperability:** A European infrastructure can address the challenges of data fragmentation and interoperability of data, models, and tools that hinder VHT development. Establishing common standards and protocols for data sharing and model development ensures seamless data integration across national borders and facilitates the creation of comprehensive VHT models.
- **Credibility and Trust:** A centralized European infrastructure can play a vital role in establishing the credibility and trustworthiness of VHT technologies. Implementing robust validation and verification processes, adhering to ethical guidelines and involving regulatory bodies ensures that VHT models are accurate, reliable and meet the highest scientific and ethical standards.
- **Economic Benefits:** A thriving VHT ecosystem can drive economic growth and create new opportunities for European industry and innovation. By fostering collaboration between academia, industry, and healthcare providers, the VHT initiative can stimulate the development of new technologies, products, and services in the healthcare sector, enhancing Europe's global competitiveness.

5.2 The current status of the VHT

While integrated Digital Twins – which bridges multiple organs and/or time and length scales – remain largely confined to the research stage, there are ample examples of the implementation of Twin technologies in Healthcare across various medical fields, focusing on specific organs, diseases or clinical applications. A comprehensive overview of all available DTs is beyond the scope of this roadmap. The figure below shows examples of existing DTs, while the subsequent Use Case descriptions provide a high level description of VHT applications that will be used throughout this roadmap to illustrate the different elements related the development and uptake of DTs in healthcare.

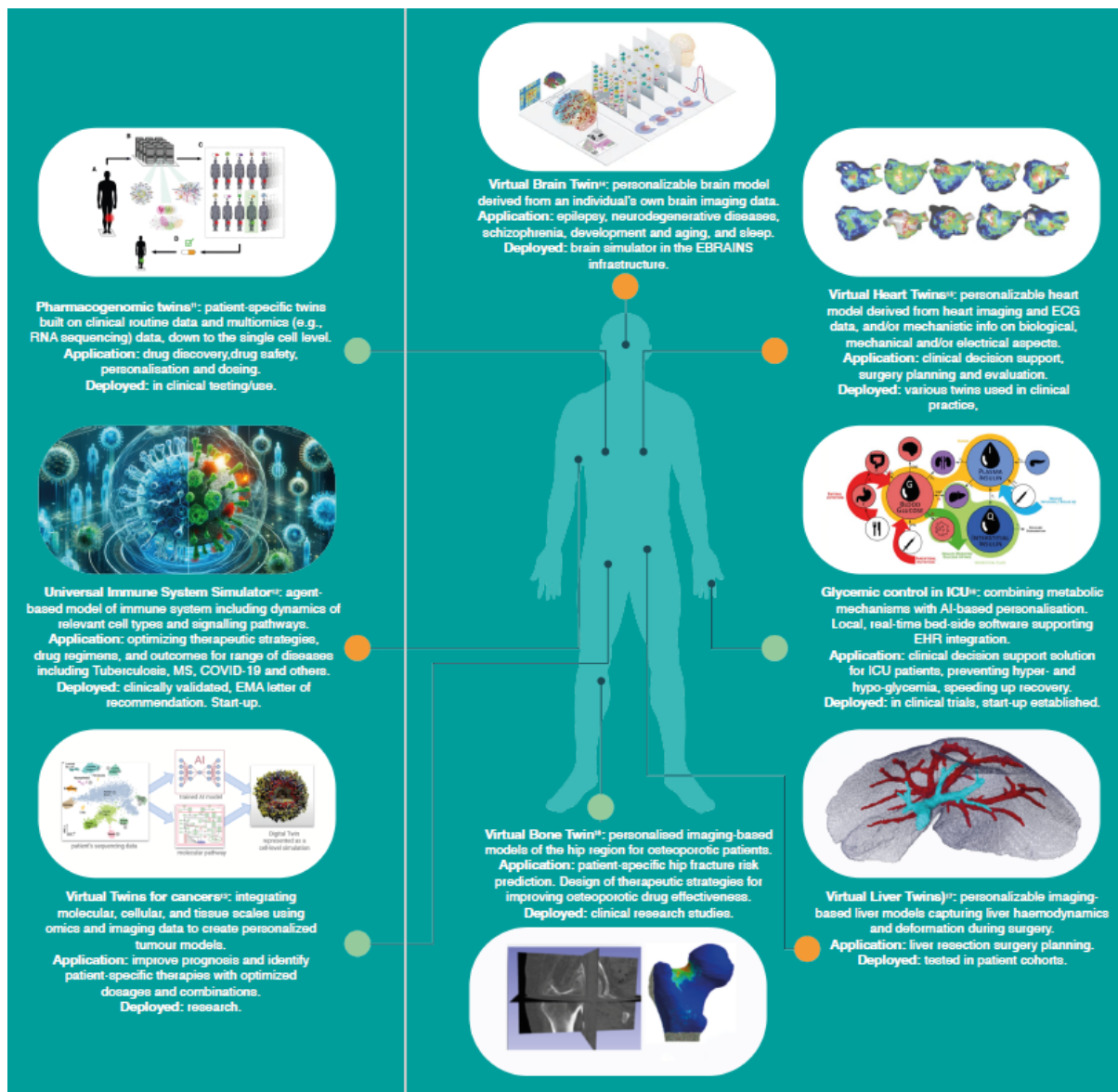


Figure 2: VHT state of the art. These examples are a few of many developments at different Technology Readiness Levels, going from basic research to clinical uptake. Further information on the examples shown here can be found in references²⁹.

Box 2: Use Case - Glycemic control in ICU patients (context)

Use Case: Glycemic control in ICU patients

Status: start-up company, solution in clinical trial

Website: www.insilicare.com

²⁹ Immune Simulator: <https://www.combine-group.org/>;

Cancer twin: https://permedcoe.eu/wp-content/uploads/2023/03/PerMedCoE_2ndPR_2023.pdf;

Brain twin: <https://www.ebrains.eu/news-and-events/using-ebrains-modelling-tools-to-investigate-the-relationship-between-brain-structure-and-function-2>;

Heart twin: <https://www.sems.qmul.ac.uk/staff/c.roney/research/>;

Glycemic twin: Uyttendaele et al. BioMed Eng OnLine 18:102, 2019. <https://doi.org/10.1186/s12938-019-0720-8>. Copyright CC BY 4.0;

Liver twin: <https://team.inria.fr/simbiotx/>;

Bone twin: Aldieri et al. Comput Methods Programs Biomed. 240:107727, 2023. <http://doi.org/10.1016/j.cmpb.2023.107727>. Copyright CC BY-NC-ND 4.0

Approximately 30-50% of all intensive care unit (ICU) patients experience glucose balance impairment, a **major factor in hospital costs, morbidity and mortality**. Glycemic control (GC) in ICU patients using insulin therapy demonstrated clinical benefits but is difficult to achieve safely due to high metabolic variability, increasing the risk of hypoglycemia.

The potential benefits of GC are often undermined by protocols that fail to account for patient-specific variability, making it challenging to achieve safe and effective control across diverse patient populations. InSiliCare's AI-powered Digital Twin offers a clinically validated solution that personalizes treatment, optimizes safety, and reduces ICU costs for all patients.

The AI-powered Digital Twin is a clinical decision support solution for ICU patients requiring glycemic control. It provides personalized, safe, and effective management of blood glucose levels and nutrition delivery. Insulin and nutrition treatments are calculated to maximise safety from hypoglycemia, while controlling patient blood glucose levels and optimizing nutrition towards a configurable physician-determined practice of care. Designed for use in adult ICU patients with hyperglycemia, the solution is not a substitute for, but rather an adjunct to clinical reasoning. The measurements and calculations generated are intended to be used by qualified and trained medical personnel in evaluating patient conditions in conjunction with clinical history, symptoms, and other diagnostic measurements, as well as the medical professional's clinical judgment. No medical decision should be based solely on the recommended guidance provided by this software program.

Box 3: Use Case - Osteoporotic fracture risk prediction (context)

Use Case : *Bologna Biomechanical Computed Tomography for osteoporotic fracture risk prediction*
Status: *clinical research studies*

Hip fractures in osteoporotic patients significantly impact quality of life, morbidity, and mortality while also imposing substantial costs on healthcare systems. Enhancing the accuracy of patient-specific fracture risk prediction can enable more effective preventive measures, reducing fracture incidence along with its associated social and economic burdens. The central clinical question a doctor must address when prescribing the most appropriate treatment is: Is the patient at high risk of osteoporotic fractures?

Bologna Biomechanical Computed Tomography (BBCT) is a Digital Twin methodology designed to predict the mechanical strength of the femur under critical loading conditions in osteoporotic patients. Quantitative Computed Tomography (QCT) scans of the hip region and patient data inform a subject-specific Finite Element (FE) model able to predict the risk of hip fracture at the time the CT is performed (ARF0).

BBCT assists clinicians in identifying patients at high risk of femur fractures, enabling the timely initiation of necessary treatments and effective osteoporosis monitoring. In addition to model results, other relevant evidence that can aid in addressing the clinical question includes bone mineral density (BMD) data, health and lifestyle factors, and information on treatment effectiveness.

Box 4: Use Case - Universal Immune System Simulator for Tuberculosis (context)

Use Case : *Universal Immune System Simulator for Tuberculosis*

Status: *augmenting clinical trials*

Website: <https://www.mimesis.srl/>

Tuberculosis (TB) remains a major global health challenge, with millions of new cases and fatalities each year. Effective treatment of TB requires precise modelling of immune system responses to the pathogen *Mycobacterium tuberculosis* (Mtb). The Universal Immune System Simulator (UISS) offers a robust, clinically validated platform to address these challenges by simulating and predicting immune responses *in silico*. Developed within the EU-funded *In silico* World (ISW) project, UISS-TB has demonstrated its utility in optimizing therapeutic strategies, evaluating drug regimens, and predicting outcomes. Its development is further validated by a Letter of Support from the European Medicines Agency (EMA)³⁰, underscoring its relevance for regulatory acceptance and clinical adoption.

UISS-TB-DR is designed to predict the dynamics of circulating interferon gamma (IFN- γ) levels over time as a function of treatment dose in virtual patient cohorts. Its primary use is to support dose regimen selection in escalating dose phase IIa trials of new therapeutic vaccines targeting pulmonary TB. This innovative approach enables the simulation of *in silico*-augmented clinical trials, significantly increasing the numerosity and variability of studied populations compared to traditional methods.

³⁰ https://www.ema.europa.eu/en/documents/other/letter-support-universal-immune-system-simulator-tuberculosis-disease-model-uiss-tb-dr_en.pdf

Specifically, UISS-TB-DR integrates patient-specific inputs, such as immune response profiles, preclinical data, and vaccine characteristics, to predict dose-response relationships. The outputs generated include dose-response curves and recommendations for intermediate doses to test alongside the minimum effective dose (MED) and maximum tolerated dose (MTD).

By enabling comprehensive simulations across a wide range of doses and virtual patient profiles, UISS-TB-DR enhances the reliability of dose selection processes while reducing the dependence on small physical cohorts. The platform supports clinical decision-making by complementing data from experimental studies, providing robust insights into the immunogenicity of vaccine candidates.

UISS-TB-DR is not a substitute for clinical judgment but rather an adjunct to enhance decision-making. Its outputs, including immune response predictions, are intended for interpretation by trained professionals alongside clinical trial data and patient history.

Box 5: Use Case - Epileptogenic zone localisation for surgical planning in epilepsy patients (context)

Use Case: Epileptogenic zone localisation for surgical planning in epilepsy patients

Status: in clinical use

Website: <https://www.cloudsoftcare.com/>

Approximately 30% of patients with epilepsy do not respond to medication and are diagnosed with refractory epilepsy. For these patients, epilepsy surgery is often the most effective treatment to achieve seizure freedom. The primary need is to precisely identify the epileptogenic zone—the brain area where seizures originate. During the presurgical planning phase, a multidisciplinary team assesses whether surgery is a viable option. This process is highly complex and involves multiple clinical evaluations, including long-term EEG monitoring and structural MRI scans. Analyzing these data is time-consuming and subjective, as it was traditionally performed through visual inspection by expert clinicians.

Persyst ESI powered by Epilog³¹ is a neuroimaging solution that automatically combines scalp EEG data with a patient's MRI to perform Electrical Source Imaging (ESI). It pinpoints the origin of brain activity linked to seizures, helping clinicians accurately localize the epileptogenic zone—critical for surgical planning in epilepsy patients.

- **Diagnostic aid:** ESI provides critical insights into seizure localization, supplementing conventional imaging techniques like MRI, PET, or SPECT. It is particularly valuable for patients with epilepsy who are candidates for surgery, ensuring precise targeting of the epileptogenic zone.
- **Practicality:** the solution integrates seamlessly into existing clinical workflows, requiring no additional hardware and supporting standard EEG setups, making it accessible and efficient for routine use.

By offering detailed, non-invasive diagnostic information, Persyst ESI improves the accuracy of epilepsy evaluations and enhances patient outcomes.

Box 6: Use Case - the Atrial Modelling Toolkit for cardiovascular Digital Twins (context)

Use case : the Atrial Modelling Toolkit for cardiovascular Digital Twins

Status: for research purposes

Website: <https://github.com/pcmlab/atrialmtk>

Atrial fibrillation (AF) is the most common cardiac arrhythmia, affecting over 37 million people worldwide. AF is associated with increased risk of cardiovascular diseases, stroke, and death. Treatment for AF includes anti-arrhythmic drugs and catheter ablation therapy. AF patients represent a diverse population, with large variations in atrial anatomy and electrical properties. These properties affect arrhythmia mechanisms and outcome of different catheter ablation and anti-arrhythmic drug treatment approaches. Virtual cohorts capture variability in atrial anatomy, fibrotic remodelling, and electrical properties across a population, and may be used to simulate the effects of different treatment approaches. This has the potential to inform personalised therapy approaches, reduce AF recurrence, reduce AF burden, and decrease healthcare costs.

To enable large *in silico* trials and personalised model predictions on clinical timescales, it is imperative that models can be constructed quickly and reproducibly. The Atrial Modelling Toolkit³² (**atrialmtk**) aims to overcome the challenges of constructing cardiac models at scale through the development of a robust, open-source pipeline for bilayer and volumetric meshes for atrial models. The atrialmtk toolkit takes segmentation masks of the atria, which can be produced from raw MRI or CT data by an expert or input directly where the

³¹ <https://www.persyst.com/technology/electrical-source-imaging/>

³² <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2023.0038>

user already has them, or existing atrial surface meshes as input. After the identification of some key anatomical landmarks, the model automatically produces a simulation grade atrial mesh that incorporates atrial regions and atrial fibres, which are input into an electrophysiological simulation. The output of this simulation can be used to investigate the effects of fibres and fibrosis on fibrillatory dynamics in the atria and to trial different treatment approaches.

This atrial Digital Twin platform **could be used as a resource by many communities** including: computer scientists who require meshes for developing simulation platforms or deep learning algorithms; by biomedical engineering researchers for running mechanistic biophysical simulations; by basic scientists who are interested in the progression of disease over time; by cardiac electrophysiologists, research academics at the US Food and Drug Administration or in other industry for simulating the response to novel atrial fibrillation catheter ablation or drug therapies as a virtual cohort trial. These models could also be used to test signal processing algorithms by researchers or electroanatomic mapping companies, or to generate statistical shape atlases. Depending on data ethics, models can be saved alongside demographics making it possible to test responses by age, biological sex, or BMI. Anatomical models can also be synthesised to represent different ages, sex or BMI through statistical shape or generative AI approaches.

5.3 A glance into the future of the VHT

To illustrate what possible outcomes could be expected from the VHT initiative, several hypothetical user stories have been elaborated for different end users³³. The scientific and computational elements that are mentioned in these stories may be already mature in some cases, while in others they may still be in a research or conceptual phase. Key elements in these stories that currently are not available are the existence of a consolidated VHT infrastructure for generating improved scientific or commercial outcomes, as well as the seamless integration of VHT applications in the clinical workflow to deliver improved clinical outcomes. Such integration of all the tools and technologies, along with the required engagement and cooperation from all relevant stakeholders, is essential to realise the vision of the VHT. *It is important to note* that these stories are merely examples of what the VHT could help to realize. They are not the only possibilities or applications, nor should they be interpreted as a statement of priorities.

5.3.1 User story: patient perspective – ultimate VHT beneficiary

Box 7: User story – a patient's journey with digital twins in healthcare

User story: a patient's journey

Mrs. Rosa (55) is a dynamic lady with an active social life and works in the retail banking industry in a rural area of Spain. Having entered menopause a decade ago and with a family history of type 2 diabetes, Rosa is keen to maintain an active lifestyle through regular physical activity. Here, she reflects on events from a little over a year ago.

“How my digital self keeps me on track with health – amidst the chaos”.

During a regular follow up with my family doctor, Dr. José, I was excited to tell him about my fitness progress. I was especially keen in sharing my steps count and the 60k step challenge I was doing with my colleagues. Dr. José was impressed and suggested I use the **VHT platform** with data pods to keep track of my health data and medical records. He explained that it's like a digital health closet that securely stores everything in a personal data pod. This sounded like the safety lockers in my bank, albeit digital and customised for health data. He assured me it follows all the EU regulations, so my privacy and security are guaranteed.

After checking out VHT, I decided to upload my smartwatch data and medical reports to my Spanish data pod. Dr. José also recommended the **Bone-Twin program**, a digital twin for musculoskeletal health, because of my age and family history of type 2 diabetes. With my consent, Bone-Twin used my activity data to assess my skeletal health and check for potential risks like osteoporosis. Thankfully, the digital twin informed me that my bone health was fine and would continue to monitor for any anomalies.

The Bone-Twin program also suggested I try the **Diabetic-Twin program** due to my family history. I was a bit hesitant at first because I felt my diet and regular blood tests were enough to manage my diabetes risk.

³³The clinician's and patient's perspective stories are inspired by scenarios described in the Discipulus roadmap (October 2013). <https://www.vph-institute.org/discipulus.html>

However, since I was starting to feel good about using these digital health tools, I decided to share my blood glucose data with the Diabetic-Twin program.

VHT digital twins facilitates early detection and intervention

Over the next year, I noticed how my VHT digital twin programs interacted with primary care services, and provided me with simple summaries to understand laboratory results. It felt like healthcare was starting to become a bit accessible like online banking. However, as summer approached, I started experiencing muscle cramps and felt less motivated to exercise. This drop in activity triggered a warning from the Bone-Twin program, which seemed to be getting information from the Diabetic-Twin program as well. Apparently, Dr. José received the warning and asked me to come in for a check-up.

Although I was annoyed about the extra appointment, I appreciated how the clinic accommodated my busy work schedule. On arrival at the primary care, Dr. José's team physically examined me. Together we then reviewed the digital twin data, which revealed a more comprehensive picture of my health. Apparently, the Diabetic-Twin program, based on my glucose measurements and dietary habits over the past months, seemed to have detected early signs of diabetes that could be increasing my risk of osteoporosis. Dr. José explained to me that considering my age, gender, and menopausal status, the Diabetic-Twin flagged the Bone-Twin, which became extra vigilant in monitoring my activity patterns and bone health. By the time I started feeling discomfort from the cramps, the Bone-Twin had already detected my reduced mobility and sent an alert to Dr. José.

Initially, I was frustrated and concerned that my family's diabetic history was surfacing. Discussing it further with Dr. José, I realized that my proactive approach to health management, along with the insights from my care team and the VHT digital twins, allowed for early detection. More importantly, VHT also has helped Dr. José to refer me now to the osteoporosis clinic for further assessment, thereby starting a timely intervention.

Digital self empowers patients and improves care coordination

At the bone clinic, they conducted further tests and confirmed that I had an increased risk of osteoporotic fractures. Further, using my Bone-Twin, they compared different treatment options, while also weighing the impact of newly confirmed diabetic condition and dietary habits. The possible outcomes and the forecasts based on the visualizations from VHT helped me understand how medications, physical activity and dietary changes could improve my skeletal health. Thus, giving me an outlook to manage the progression of osteoporosis, while also keeping an eye on the diabetes.

Over the following year, I received multiple communications from the VHT-connected digital twin services, which, along with my doctor's consultations, empowered me to take control of my health. I was also occasionally contacted by researchers using anonymized data from my data pod for clinical trials. It felt good knowing how my data was being used and that I was contributing a bit to citizen science. It was interesting to note that one such research team in Athens, Greece, was studying the effectiveness of osteoporosis drugs in people with type 2 diabetes, which resonated with my own situation.

As part of my osteoporosis treatment, I joined an aqua-gym class, which provided the perfect balance of physical activity and social interaction. And with the help of my digital twin, I even discovered that the pool's cafeteria had some healthy snack options. I felt more informed and continued to play an active role in managing my health. During our neighbourhood's annual flu shot campaign, Dr. José encouraged me to share my experience of managing multiple health conditions with other community members at the primary care facility. Naturally, along with Dr. José's team, I couldn't leave out the positive impact of digital twins and how they've helped me, with early detection and intervention of my newly diagnosed chronic conditions.

5.3.2 User story: clinician & healthcare provider perspective – VHT user and beneficiary

Let's read through the perspective of Dr. Anna Schmidt, cardiologist at Trinity Hospital, Berlin.

Box 8: User story – a clinician's perspective.

At Trinity hospital in Berlin, the cardiology team led by senior Cardiologist Dr. Anna Schmidt, regularly receives referrals of patients experiencing non-specific symptoms, such as chest heaviness, along with potential risk factors. Unlike emergency admissions, these patients are referred not only due to recent discomfort but also because of their increased risk of cardiovascular disease, considering factors like age, family history, and genetic predisposition.

Mr. Mueller (58), a desk worker who enjoys gardening, was one such referral to the **Trinity Hospital**. He experiences chest heaviness during strenuous gardening. Despite being on blood pressure medication, his General physician (GP), considered Mueller's family history – where both parents died of heart disease in their early 60s – referred him to Dr. Anna Schmidt, at the cardiology unit at Trinity Hospital. This is her recounting of their interactions.

“Upon Mr Mueller's arrival at Trinity, my team and I conducted a thorough assessment, including physical examination, blood tests, electrophysiology assessments and imaging of his heart. Considering his family history and the recent advancements in our understanding of the genetic factors involved in complications of coronary interventions (e.g. thrombosis), we also explored the possibility of genomic testing to identify variants that could affect platelet function and increase his risk of thrombosis. The initial results confirmed the suspicions of the GP: Mr. Mueller's discomfort was likely heart-related. To gain a comprehensive understanding of his condition and explore potential treatment options, we created a **personalised digital twin of his heart** using **Trinity's Heart-Twin technology**.

The Diagnosis – Early & Accurate

The innovative Heart-Twin Technology at Trinity rapidly integrated recent clinical and non-clinical data with Mr Mueller's medical history, enabling us to diagnose the developing coronary artery disease. We discovered that multiple vessels supplying blood to his heart were affected, restricting blood flow. The **Heart-Twin's diagnostic module** provided non-invasive measurements of blockage severity in each vessel, giving us detailed insights into the blockages, their impact on blood flow, and the overall severity of the disease per vessel. The digital twin pinpointed areas of Mr Mueller's heart tissue receiving insufficient blood, highlighting a risk of infarction and a potential heart attack. This level of detail is often not achievable with traditional diagnostic methods alone. More importantly, it saves over 20 million Euros per year in emergency admissions within Germany and is fully reimbursed by the health insurance.

With these initial findings, I met with Mr. Mueller to explain the situation about his heart. Understandably, he was concerned that his health might follow the same path as his parents. Using his own digital twin representation, we showed him how we had captured the anomaly in his heart at a very early stage. I assured him that our team would place him in telemonitoring and follow up in the coming weeks with detailed analysis and personalised treatment options based on his **Heart-Twin data**.

Initially disappointed by his diagnosis, Mr. Mueller listened to our explanation and went home with the knowledge that the digital twin could assist both our care team and him, in closely monitoring and addressing his diagnosed condition.

The Prognosis and Treatment Options - Personalised

Using the **Heart-Twin's prognostic module**, we analysed potential treatment outcomes, considering his medical records, family history, and current diagnosis. We asked the Twin to simulate various treatment approaches on his digital heart, experimenting with different medications, stents, and lifestyle modifications. Crucially, the Heart-Twin model, incorporating Mr. Mueller's genetic traits, flagged a rare genetic deficiency, which could interfere with the metabolism of standard anti-clotting medications. With such comprehensive assessment and virtual simulations, the prognostic module helped us evaluate multiple treatment options as well as assess the risk of thrombosis. This helped us identify necessary precautions to ensure the optimal outcome for Mr. Mueller.

The Treatment and its planning - Precision

Our cardiac unit at Trinity, comprising cardiologists, interventional specialists, imaging specialists, biomedical engineering experts, and related professionals, convened to discuss findings from the **Heart-Twin prognostic module**. Collaboratively analysing Mr. Mueller's digital twin, we devised a **personalised treatment plan**. This plan included surgical intervention to place a drug eluting stent (limiting restenosis) to address the major blockage, followed by oral medications to dilate smaller blockages. Our clinical team deliberated about the Heart-Twin's moderate risk score for stent thrombosis, based on genetic profiling. However, no further treatment adjustments were agreed upon, except for substituting the standard anti-clotting drug with an alternative due to Mr. Mueller's genetic deficiency in metabolising the default drug.

Subsequently, our cardiac surgeons tasked the **Heart-Twin's planning module** to simulate the surgical procedure on Mr. Mueller's digital heart. After considering a range of optimal access points, ideal stent locations, and selecting the most suitable implant devices based on his individual anatomy and needs, the planning module generated a clinical report. The module also contains the option of using extended reality technology (AR/VR) to visualise the surgery in case certain complexities are expected.

The Patient Engagement – Informed & Active role

Upon Mr. Mueller's return to Trinity in the following weeks, we first acknowledged his apprehension regarding the diagnosis. Then, our care team invited him to don VR goggles. This provided an immersive visualisation of the procedures and potential outcomes of different treatment options using the Heart-Twin program, empowering him to make an informed decision about his care. In addition, we also walked him through the digital twin's predicted outcomes for the coming years, incorporating factors like his medication history, various lifestyle scenarios (cholesterol levels, blood pressure, activity levels, stress levels), and the impact of

personalised medications that counteracted his genetic predisposition. These predictions helped us to powerfully underscore the critical role of lifestyle modifications and strict adherence to medical advice in preventing the progression of vessel disease and the need for future re-interventions.

The Health Outcome – Improved

Within six weeks, Mr. Mueller underwent a successful surgical procedure, meticulously guided by the virtual planning conducted using his Heart-Twin. We then updated the Heart-Twin with post-operative images of Mr Mueller's rejuvenated heart along with the vitals collected through remote monitoring during his rehabilitation phase.

During the post-operative consultation, Mr. Mueller expressed a keen interest in how his Heart-Twin was gauging his recovery. Our cardiology team at Trinity were delighted to witness Mr. Mueller's active engagement with his digital representation and his enthusiastic commitment to the collaboratively envisioned care pathway. By leveraging the Heart-Twin, we were able to reassure him of his diligent adherence to medications and his embrace of a healthier lifestyle. This newfound confidence empowered him to resume his regular routines, including his beloved hobby of gardening.

To conclude, the above benefits empower us as healthcare professionals as well as our patients. Overall, it exemplifies how the Heart-Twin reinforces the high-quality care we strive to provide. We increasingly value the Heart-Twin and its crucial role in supporting us to deliver the best possible care and also promote healthy living for Mr. Mueller and all our patients, at Trinity Hospital, here in Berlin, Germany."

Leveraging the Heart-Twin Technology at Trinity Hospital, Berlin

For Dr. Anna Schmidt, the healthcare professional as the user of the virtual human twin: The Heart-Twin technology has proven invaluable in Mr Mueller's care. It streamlined vast amounts of clinical data and his personal medical history, facilitating personalised treatment plans, forecasts, and strategies for preventing future complications. It also addressed both clinical and non-clinical factors by accessing patient information collected at home and at the hospital, and allowed to get a comprehensive picture compared to snapshot information. More importantly, it allowed us to dedicate more quality time to patient care. Dr Schimdt and her team were able to use the Heart-Twin to simulate various treatment approaches, including medication, stenting, and lifestyle modifications. This allowed them to assess the potential outcomes of each option and develop a treatment plan tailored to Mr. Mueller's individual needs and risks. The visualisation and prognostic modules within the Heart-Twin program have been instrumental in facilitating collaboration and communication among their interdisciplinary team and with the patient, as well as better anticipate complications.

For Mr. Mueller, as the ultimate beneficiary, the patient: Early detection and intervention, lead to a better prognosis and reduced risk of serious complications and establishing a personalised treatment plan that addressed his specific needs and genetic predispositions, minimising potential side effects and complications. Active involvement in his care, thanks to the visualisations and explanations provided by the Heart-Twin, empowered him to make informed decisions about his health, track his adherence to treatment plans and promote healthy lifestyle choices. Overall, digital twin technology would also support long term proactive monitoring of heart patients like Mr. Mueller, who are at three times higher risk of comorbidities like diabetes, and are also at risk of adverse cardiac events.

5.3.3 User story: researcher perspective – VHT user, developer and beneficiary

Let's read through Julia's personal experience of using the VHT platform to advance her research on addressing the fast-evolving dual endemic of osteoporosis and diabetes.

Box 9: User story – a researcher's perspective.

Julia is a **graduate student at the University** of Bologna, where she is working on her PhD in Biomedical Engineering. Here, she talks about her recent experience with the European VHT infrastructure.

My name is Julia, and I'm a 26 year old PhD student working on a project investigating how patients with diabetes mellitus type II respond to anti-resorptive treatments, a class of drugs that slow down the progression of osteoporosis. My project involves conducting a clinical study on this topic. While the primary clinical endpoint in this type of studies would be the bone fracture event, we use the absolute risk of fracture provided by a Digital Twin called BBCT as a surrogate biomarker of this endpoint. A properly calibrated computed tomography (CT) of the thigh region is the primary input for BBCT.

Generating a VHT application – putting together the right pieces of the puzzle

A little while ago, I was contacted by another hospital who would like to join our multicentric clinical study with a cohort of patients who had already been enrolled in a previous study and that are still being followed up. The problem is that the CT scans were collected for these patients, but no calibration was available. Searching the literature, I found a phantom-less calibration method that could be used to calibrate CT scans a posteriori. But the question remains: how does this change in the BBCT protocol affect the predictive accuracy of the Digital Twin? My supervisor told me that ten years ago, when he was a post-doc in the team that developed BBCT, such a task would have required months of work. But not today: today, there is the VHT. BBCT is one of the thousands of Digital Twins available on the VHT. So, I searched for a properly calibrated CT scan of the thigh region: the HD_Valid collection, also available on the VHT, provides 100 of them. I wrote a script that implements the phantom-less calibration described in that paper, and I created on the VHT a copy of the CT scan calibrated with the phantom-less method. Then I scripted VHT to run twice: once with the phantom-less calibrated CT, and once with the properly calibrated CT. Last, I linked my script to another one already available in the VHT, which compares two quantities and calculates all sorts of error indicators such as root mean square error, peak error, etc.

Analysing the obtained results – one more iteration

I tested my script on the single CT, and it worked fine. So I executed my script on the whole VHT data space. The VHT's resource finder found 579 properly calibrated CT scans of the thigh region, shared on VHT by several researchers. I chose as execution space the Italian National Centre for HPC, where my group has a grant for several core-hours. And then I went out with my friends for aperitifs. The day after, I was informed the solution to my problem was ready. Upon accessing it through my VHT dashboard, I was disappointed. The phantom-less calibration would introduce errors in the BBCT prediction of 10% on average but with peaks of 35% for some cases. Not good.

With a little help from my VHT friends

As my script was executed, all researchers who shared the CTs I used were informed of such use and the purpose of my study. The day after, I was contacted by one of them, named Yüzé, who is working on a new methods of phantom-less calibration for his PhD at the university of Liège in Belgium. Since my validation script is not part of the VHT, it took him five minutes to replace the calibration script I implemented from that paper with his own and run the whole thing again. And this time, the peak error is only 7% across all cases! Looks like I have the solution to my problem – and a new scientific collaborator. Soon after Yüzé and I submitted a manuscript to IEEE Journal of Health Informatics, where we presented his new calibration method and my validation protocol; the paper links to the VHT resources we used for this study.

Today I submitted to the ethical committee the request for an extension for the multicentric clinical study to include the new hospital; I included the results of our *in silico* study to demonstrate that the change in the protocol has a minor effect.

5.3.4 User story: industry perspective – VHT infrastructure accelerates drug development

This fictional story illustrates the potential of the VHT infrastructure to empower companies like ACME in providing enhanced consulting services.

Box 10: User story – a company's perspective.

Dr. Przemko Kowalski, CEO of ACME Consulting Services, leads the company with over 20 years of experience in medicinal product development. Headquartered in Krakow, Poland, ACME has become a go-to advisor for Biotech and Pharma manufacturers, particularly those pursuing non-conventional development strategies to de-risk product development and bring innovative health technologies to patients. Here, Dr. Kowalski shares his story on how a robust VHT infrastructure empowers ACME's mission to address the global challenges faced by the health care players, including the World Health Organization (WHO).

A digital twin approach to drug discovery

As the CEO of ACME Consulting Services, I've spent over 20 years in medicinal product development. Our specialty is using a cutting-edge digital twin of the human immune system. This allows us to test new drugs or repurpose old medicinal products *in silico* before any animal or human experimentation is done. This has drastically reduced the attrition rate of drug candidates and has allowed us to solidify our unique market position, across the globe.

Confronting a new pandemic

Recently, the WHO approached us regarding a new pandemic that was starting to spread globally. This concerned a new coronavirus variant, which was causing adverse immune reactions and a hyperinflammatory

response, leading to critical respiratory distress and, in some cases, death. The WHO urgently needed to find the best available Nonsteroidal anti-inflammatory drugs (NSAIDs) to treat the respiratory symptoms while a vaccine was being developed.

Leveraging molecular data from the spike protein obtained through Cryo-electron microscopy, we rapidly developed an *in silico* trial core by extending our base immune system model (physiological layer) with a disease model of the coronavirus infection in the lungs (pathological layer). Through this, our model incorporated molecular and cellular information alongside clinical knowledge from other coronavirus-induced diseases such as SARS, MERS, and COVID-19.

Leveraging the VHT infrastructure: A Game Changer

We faced a major challenge: there were 35 different NSAID molecules available on the market, each working differently. Creating 35 individual intervention models is an incredibly time-consuming and resource intensive exercise! Luckily, the EU-driven VHT infrastructure came to our rescue. The VHT repository had pre-existing interventional models for 23 of the 35 treatments. As strong believers in the power of collaborative public-private research, we had designed our simulator to be fully integrated and compatible with the VHT infrastructure. This allowed us to easily integrate those 23 interventions into our *in silico* trial.

Simultaneously, our ACME team was building a virtual patient group for the *in silico* trial. Utilizing the epidemiological information provided by WHO and collating our insights from the COVID-19 pandemic, my colleagues aimed to create a virtual cohort representing the at-risk patient population. However, even with cohort expansion techniques, collecting experimental data for a few hundred subjects within the target population would take months. The VHT repository again came to our rescue. Our teams noted that the VHT platform hosted two large datasets of COVID-19 patients. Using this data and the VHT's phenotype interpolation models, we built a virtual cohort of 3000 patients in just a few days.

Accelerated response: a new era in pandemic preparedness

Two weeks after the project started, working days and nights, we were able to launch the first *in silico* trial. Given the importance of the health threat, EuroHPC made a special allocation to our trial, dedicating half of its tier-0 supercomputer. In a matter of days, one specific candidate treatment emerged, among the 23 tested. WHO initiated clinical trials worldwide, and all clinical data were collected on the VHT infrastructure, where the automated credibility assessment pipelines were used to validate the predictions of the ACME model.

Low predictive accuracy was found for a specific subpopulation; the problem is that, for two NSAID molecules in a subgroup of patients who regularly take another drug, interference between the two treatments arises. Testing such interference in an ad-hoc *in silico* trial, confirmed the hypothesis. A week later, WHO experts were able to provide guidelines on handling this specific sub-group.

Thanks to the VHT infrastructure and its vast collection of resources, we were able to identify the first line of treatment only two months after the first case was reported. This is a drastic reduction compared to the seven months that it took in 2020 to identify dexamethasone as an optimal treatment for critically ill patients infected with COVID-19. With all our data and models now on the VHT infrastructure, we're even better prepared for future pandemics.

6 The Virtual Human Twin Ecosystem

Realising the vision of the VHT is beyond the capabilities of any individual organization or even community (*e.g. in silico* medicine, wearables, organ-on-chip), but rather requires establishing and **engaging the entire ecosystem**. The concept of ecosystems encompasses the intricate interconnections between various sectors and stakeholders. The VHT ecosystem includes all participants along the VHT's value chain, including academic and research organisations, industrial organisations from the smallest start-up to the largest corporations, healthcare professionals, patients, regulators, health authorities, payers, policymakers, and more. The figure below provides an overview of the entire VHT Community of Practice (CoP), starting from the *VHT beneficiaries and users* identifying the need, over *VHT creators* developing solutions and *VHT enablers* providing the necessary infrastructure and resources, to the *VHT managers* responsible for its governance.

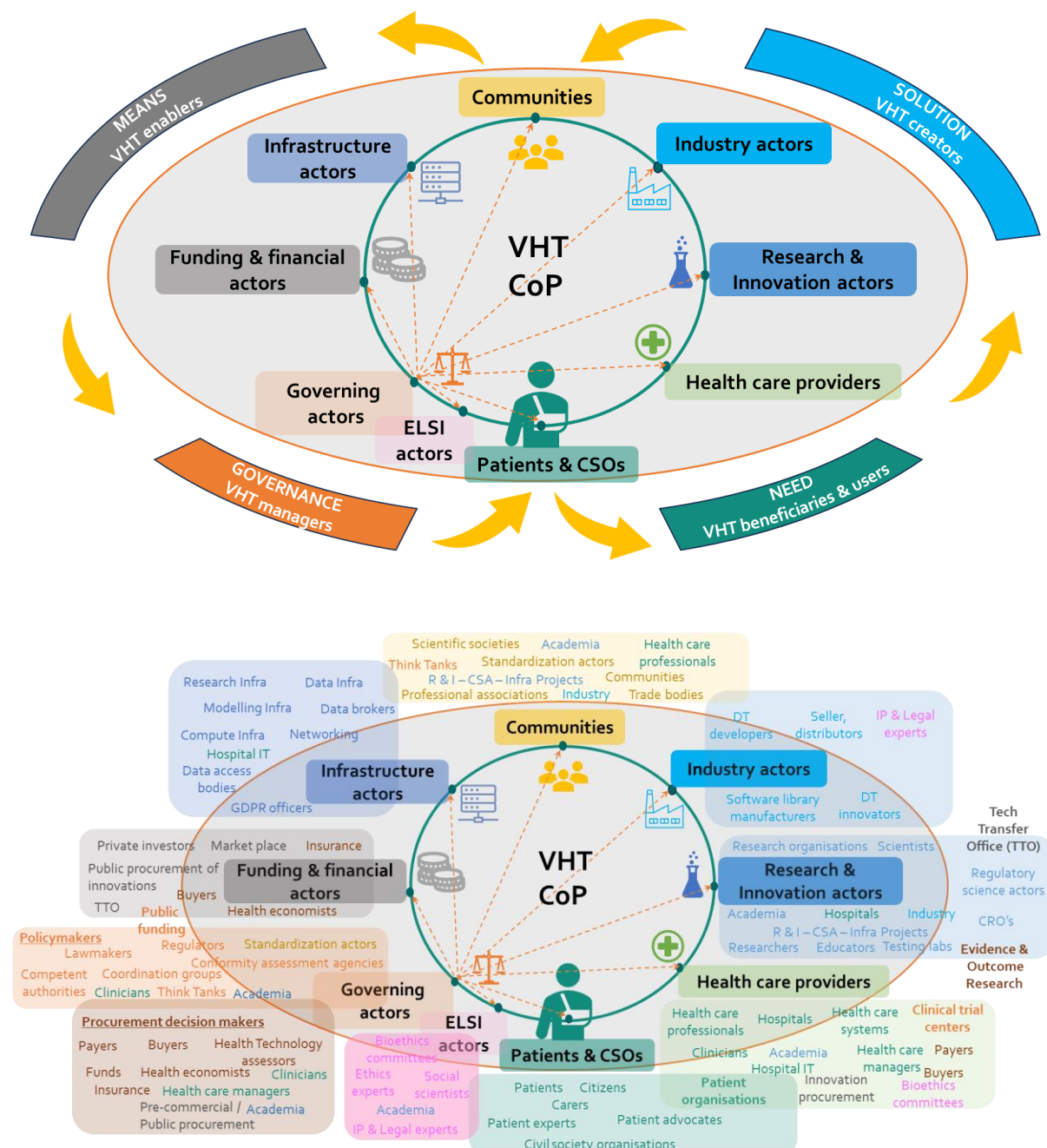


Figure 3: Schematic depiction of the VHT Community of Practice (CoP) and its different stakeholders. ELSI: Ethical, Legal, Social Issues; CSO: civil society organisation.

6.1 VHT beneficiaries and users

6.1.1 Patients & Citizens

Patients and citizens are **beneficiaries and users** of VHT technologies, who are central to the VHT CoP. They are involved in co-developing and co-assessing VHT technologies. Patient and citizen groups are end users or beneficiaries of Digital Twins in Healthcare, where DTs are addressing an unmet clinical need. They can be further categorized as follows:

- **Patients** with personal experience of living with a disease, benefiting from Digital Twins to diagnose, treat, manage or predict long term evolution of ailments.
- **Citizens** are common people like patients, representing the broader society and likely to use VHT for accessing their own Digital Twins and personal medical data to better understand their health.
- **Carers** supporting patients as a family member, friend or volunteer, who can use Digital Twins to understand the treatment pathways and situate / support the patient's care.
- **Patient advocates** that have insights on large populations living with specific disease, may gain insight on population Digital Twins to channelize further research on treatments, or use Digital Twins for advocacy on policies, research funding and technology assessment.
- **Patient experts** with technical knowledge contributing to co-creating DTs or championing patient reported outcomes to establish the value of DTs within regulatory and reimbursement assessments.
 - Example: EUPATI trained Patient Experts, playing active roles in advisory boards, technology assessment committees, etc.
- **Patient organizations** mandated to voice the collective view of patients in the development, approval, and reimbursement of DTs that tackle a specific issue or disease.
 - Example: European Liver Patients Association (ELPA)
- **Citizen scientists** are individuals with limited health technology background and ambition to perform citizen/community driven research, who are likely to use VHT for citizen science research initiatives.
 - Example: Civil society organization - EU4Health³⁴

Patient organizations are typically not-for-profit organizations which are patient focused and where patients³⁵ represent the majority of members in the governing bodies. There are many types of patient organizations, with some focusing on the local or regional level, and others on the national, European or international levels. Some organizations or coalitions work across different diseases and aim to represent the voice of the whole patient community on cross-cutting issues. Other organizations deal with a specific disease (*e.g.*, breast cancer) or disease-area (*e.g.*, cancer). Patient organizations that focus on a specific disease and/or operate on a local or regional level are typically very close to the community they represent, with detailed knowledge on the challenges and needs. Examples of patients organizations include **ELPA**³⁶ (European Liver Patients Association) and **EMSP**³⁷ (European Multiple Sclerosis Platform).

In order to increase their impact on the global level, organizations sometimes form broader coalitions across diseases or countries. **The European Patients' Forum**³⁸ (**EPF**) is an example of such a broader organization, grouping 78 members³⁹ that are either pan-European or national organizations themselves. EPF has developed five criteria⁴⁰ to ascertain whether organizations are true patient organizations. These criteria include transparency, legitimacy, democracy, representativeness and accountability & consultation. EPF leads the patient advocacy in Europe, acting as an intermediary between the patient community and the EU policy makers, participating in stakeholder groups, and providing training to increase health literacy and patient empowerment. Another organization that operates in this space but is not a patient organization itself, is **EUPATI**⁴¹, the European Patient Academy on Therapeutic Innovation. They focus on educating patients and training patient experts that are able to get involved in relevant research and development, as well as educate researchers on how to engage with patients.

³⁴ <https://eu4health.eu/>

³⁵ Or carers in case patients are unable to represent themselves

³⁶ <https://elpa.eu/>

³⁷ <https://emsp.org/>

³⁸ <https://www.eu-patient.eu/>

³⁹ https://www.eu-patient.eu/globalassets/list-of-members_ec_2024.pdf

⁴⁰ <https://www.eu-patient.eu/Members/what-is-a-patient-organisation/types-of-patient-organisations/>

⁴¹ <https://eupati.eu/>

EPF can be considered as a **civil society organization (CSO)**, being an organization that is building and nurturing a democratic society based on fundamental rights. A wide range of CSO's has joined forces in Civil Society Europe⁴². Examples of other CSOs are the European Youth Forum⁴³ (focusing on youth rights and participation), European Civic Forum⁴⁴ (focusing on democracy, human rights and climate justice), and the Center for European Volunteering⁴⁵, which is the European network of over 60 organizations dedicated to the promotion of, and support to, volunteers and volunteering in Europe at European, national or regional level.

6.1.2 Healthcare professionals & systems

Healthcare professionals and systems are **users, beneficiaries and co-creators** of VHT technologies, with oversight on safety-risk-benefit-harms and adverse events. They are involved in identifying unmet needs, developing, assessing (clinical/regulatory/reimbursement) and adopting VHT technologies within healthcare systems (users → developers → providers → payers). They are composed of following groups:

- **Clinicians** whose needs are an important driver and who may use Digital Twins for increasing their understanding of specific cases and co-morbidity, as clinical decision support systems or as a way to visualise (patho)physiology and therapeutic approaches to students and patients.
 - Example network: Clinical societies like Biomed Alliance, European Society of Cardiology (ESC).
- **Healthcare providers** (Nurses, Physiotherapists, Psychotherapists and allied healthcare professionals) who will be working with Digital Twins to monitor patient progress and provide or adjust care based on the Digital Twin predictions.
 - Example: European Federation of Nurses Associations (EFN), European Nursing Council (ENC), European Association for Psychotherapy (EAP).
- **Healthcare providers, Hospitals and their IT departments** (*e.g.* Clinical IT Manager, Medical Researcher, Medical Educator), who will need to integrate Digital Twins in their daily operations.
 - Example: European hospital and healthcare federation (HOPE), European University Hospital Alliance (EUHA).

Clinical and medical organizations play an important role in the clinical community. They work on identifying common problems and challenges in the community, are involved in improving medical education, inform their membership of innovations and policy changes, disseminate new insights and technologies (*e.g.* through the co-creation of clinical guidelines) and are the key intermediaries in dealing with policy and regulatory agencies. Several organizations are actively looking into initiatives related to Digital Twins, such as the **European Society for Cardiology**⁴⁶, the **European Respiratory Society**⁴⁷ and the **European Society for Radiology**⁴⁸. 35 medical organizations joined forces in an umbrella organization, **BioMed Alliance**⁴⁹, through which they collaborate on research projects, the organization of continued medical education and facilitate interactions with the European Commission with a view on strengthening the funding for medical research, as well as discussing other policies related to European health research and on enhancing the visibility of EU health research. **Other healthcare providers** (nurses, physiotherapists, psychotherapists) have also formed representative societies active at national and European level (*e.g.*, European Federation of Nurses Associations⁵⁰ or the Europe Region of World Physiotherapy⁵¹), with the same goals as the clinical organizations. In addition to societies uniting healthcare providers (HCP), there are also **societies uniting hospitals** such as **HOPE**⁵² (European Hospital and Healthcare Federation) and **EUHA**⁵³ (European University Hospital Alliance). These associations aim to improve healthcare, research and education in order to

⁴² <https://civilsocietyeurope.eu/>

⁴³ <https://www.youthforum.org/>

⁴⁴ <https://civic-forum.eu/>

⁴⁵ <https://www.europeanvolunteercentre.org/>

⁴⁶ <https://www.escardio.org/>

⁴⁷ <https://www.ersnet.org/>

⁴⁸ <https://www.myesr.org/>

⁴⁹ <https://www.biomedeurope.org/>

⁵⁰ <https://efn.eu>

⁵¹ <https://www.erwcpt.eu/>

⁵² <http://www.hope.be/>

⁵³ <https://www.euhalliance.eu/>

improve patient care and outcomes and to enhance efficiency and effectiveness in the operation of hospitals and healthcare services.

6.2 VHT creators

6.2.1 Research and innovation actors

Research and innovation actors are **creators/developers** of VHT with a focus on basic research and innovation of VHT technologies, beyond state of the art. They furthermore are involved in developing regulatory and evidence science, governance framework including ELSI, liaising as expert assessors and appraiser of VHT technologies and their impact across the entire health continuum (users → developers → enablers → governance). In addition, research & innovation actors also address the education and (re-) training needs of the VHT CoP, as well as act as a catalyst through scientific societies and professional associations, where they contribute towards knowledge sharing and consensus building in the process of defining cross-community VHT related standards. Likewise, the research innovation actors fill the role of independent reference laboratories or testing centres, as well as conducting clinical research studies in collaboration with academic medical centres and industrial actors.

- **Researchers** in academia and industry, who develop new knowledge and new methodologies and test them in pre-clinical and clinical settings.
 - Digital Twin researchers are spread out in Academia, industry, research organizations. Specifically, the **academic sector** is the predominant research-oriented technology developer stakeholder group, which tackles fundamental challenges of conceptualization and builds Digital Twins, often providing the basis for other technology developer actors to experiment, scale up and operationalize the use of Digital Twins in healthcare systems.
- **Academia & research institutions** are prime drivers of innovations around Digital Twins in healthcare, who are typically supported by public and private funders, whereby they are often involved in fundamental research and knowledge needed for Digital Twins, and also play an active role in prototyping, testing, regulatory science for Digital Twins in healthcare.
- **Hospitals and health institutions** that use, test, (co-) develop in-house to commercial Digital Twins in healthcare. Together with their academic and industrial partners, they are important actors to conceptualize care pathway centric Digital Twins for clinical decision systems, initiated as a ‘research-only use’ or ‘off-label’ use in hospitals. They are often the first to spot unmet patient needs to trigger the opportunity and scope of a Digital Twin solution and also the first actor to use, apply or introduce the Digital Twins in healthcare systems.

The **Virtual Physiological Human Institute**⁵⁴ (VPHi) is an international non-profit organization, whose mission is to ensure that the Virtual Physiological Human is fully realised, universally adopted, and effectively used both in research and clinic. To this end, it organises its activities both on the scientific level and the policy-regulatory level, ensuring the entire path from computer screen to the patient is rolled out and all stakeholders are involved. The **Avicenna Alliance**⁵⁵ is an association of industry and research organizations who have a commercial or research interest in the development of *in silico* medicine. This Association bridges the gap between the scientific community, industry and policy makers by advocating for policy changes that take into account scientific and market developments. The **COMBINE**⁵⁶ (COMputational Modelling in BIOlogy NETwork) network is a grassroots initiative driven by the scientific biocomputational modelling community, that develops open community standards and formats for computational modelling. It aims at identifying gaps or addressing new needs in the systems and synthetic biology standards landscape and to develop interoperable standards with minimum overlap between them.

In addition to societies on the *in silico* elements, other societies focus on technologies related to data generation. **EUROoCS**⁵⁷ is the European society for organ-on-chip technology, one of the key technologies that is able to generate human-relevant data in the absence of clinical measurements. Large

⁵⁴ <http://www.vph-institute.org>

⁵⁵ www.avicenna-alliance.com

⁵⁶ <https://co.mbine.org>

⁵⁷ <https://euroocs.eu/>

organizations focusing on technology development and/or uptake such as **IEEE**⁵⁸ have an increasing focus on health applications, with large global societies on engineering in health such as **IEEE-EMBS**⁵⁹ (Engineering in Medicine & Biology Society).

6.2.2 Industry DT creators

Industry plays an important role in the creation/development of the VHT with a focus on applied research, development and industrialization of VHT technologies, in compliance to standards, regulations, legislative and societal guard rails. They are furthermore involved in the practice of regulatory and evidence science, governance frameworks throughout the lifecycle and ecosystem of VHT technologies. There is a wide range of industrial profiles relevant for the VHT:

- **OEMs** (Original Equipment Manufacturers) in the VHT context, refers to software developers that provide libraries and solvers used to implement the Digital Twins; they are cross-sectoral actors that empower the healthcare technology developers to harness the libraries developed and deployed in sectors like aviation, industry, automobile etc. They partner up with both academic and industrial VHT actors.
 - Multi-physics modelling & simulation software developers;
 - Example: Commercial and open-source simulation platforms like Ansys, Matlab, etc.
- **Medical product developers** who use Digital Twins to design, optimise, or assess the safety and efficacy of new medical products, both in the pre-regulatory and regulatory phases;
 - Device, medicinal or biologics product manufacturers;
 - Example: Medtronic, Siemens, Philips, etc.
- **Innovators** who translate research results into potential solutions for clinical or industrial unmet needs;
 - Start-ups, Small-Medium Enterprises (SME), Large Enterprises, multi-national companies (MNC).
- **Industrial sector** often collectively represents many of the above sub-groups of technology developers, whereby they are involved in prototyping, developing, testing, seeking regulatory approval and market access of proprietary Digital Twin components, products, platforms, solutions, systems or services in the digital health technology space. They further follow the lifecycle of the Digital Twins in healthcare, be it for mass manufacturing, deployment, maintenance and support.
- **Digital Twin sellers, distributors and marketers**, which can be categorized depending on their business model:
 - *Biomedical instrumentation sellers* that sell the Digital Twins as software embedded in their hardware;
 - *Medical device sellers* that sell Digital Twins as algorithms embedded in their medical devices or that complement them;
 - *Medical software sellers* that sell Digital Twins as stand-alone products or services;
 - *Broker sellers* that sell Digital Twins developed by third parties, usually as software as a service (SaaS), Platform as a service (PaaS) or Infrastructure as a service (IaaS).

Trade organizations represent industrial organizations in specific segments of the industry. **MedTechEurope**⁶⁰ is the European trade association and represents the diagnostics and medical device industry in Europe. **COCIR**⁶¹ (European Coordination Committee of the Radiological, Electromedical and Healthcare IT Industry) is a non-profit trade association. Among others it cooperates with ISO contributing to standards in health informatics. It is also engaged in services concerning medical regulatory frameworks and legal affairs. **EFPIA**⁶² (European Federation of Pharmaceutical Industries and Associations) is a lobbying organization representing the research-based European pharmaceutical industry promoting the use of medicines and medical products. The EFPIA Disclosure Codes are self-regulatory transparency rules for the pharmaceutical industry.

⁵⁸ <https://www.ieee.org/>

⁵⁹ <https://www.embs.org/>

⁶⁰ <https://www.medtecheurope.org/>

⁶¹ <https://www.cocir.org/>

⁶² <https://www.efpia.eu/>

6.3 VHT enablers

6.3.1 Infrastructure actors for Data, Research and Computing

Infrastructure actors for data, research and computing are VHT enablers or orchestrators who facilitate the effective design, development, implementation, deployment and maintenance of VHT technologies, in compliance to standards and regulations. They are important in the co-development of the VHT, empowering end-users and beneficiaries of the CoP to access the VHT technology in practice.

- **Research infrastructures** are facilities, resources and services developed and used by the scientific and R&D community to conduct high-level research.
 - Example: European Research Infrastructure Consortia (ERIC)
- **Computing infrastructures** provide compute services to the community. Several centres for supercomputing exist across Europe, with a number of organizations aiming to coordinate access to leading-edge computing and data management resources and services for large-scale scientific and engineering applications at the highest performance level.
 - Example: High performance computing infrastructures such as MareNostrum, Leonardo, LUMI etc; overarching organizations such as PRACE or EuroHPC.
- **Data infrastructures** in the medical context enable the health and care system to access data while safeguarding the privacy of the individuals involved. Applications for the use of these data include improvement of healthcare outcomes, efficiency of services or augmenting the impact of research.
 - Example: EOPYY(Greece), NHS Digital (UK), Sundhed (Denmark).
- **Data brokers** that curate and resell use licenses for data collections usually generated by third parties;
 - Example: Epic (USA).
- **Health Data Access Bodies (HDAB)**. As part of the European Health Data Space (EHDS), each country must establish at least one Health Data Access Body (HDAB) as a governance structure and qualified intermediary for the secondary use of health data⁶³.
 - Example: Findata (Finland), Health Data Hub (France), The Danish Health Data Authority (Denmark), Belgian Health Data Agency (Belgium).

6.3.2 Technology investors, accelerators, governmental and non-governmental funding agencies

Technology investors, accelerators, governmental and non-governmental funding agencies fund the **development of innovative VHT technologies**. They provide mentorship, coaching, and incentives to support the successful development of the first generation of VHT solutions. Their activities can be public, private or through public-private partnerships. They streamline and optimize the innovation pathway within the VHT ecosystem to effectively address unmet healthcare needs.

- **Business angels and investors** who support the creation of new companies that want to sell Digital Twins.
- **Tech transfer offices (TTOs)**, play a crucial role in identifying promising VHT innovations within academic environments and fostering their translation into spin-off companies. Given that sophisticated VHT models are likely to originate from academic research groups, universities need to establish dedicated programs and foster a culture of innovation to incubate early-stage VHT solutions, guiding them through the challenges of regulatory hurdles and market access.
- **Innovation hubs** complement the work of TTOs by providing a supportive environment for early-stage startups to grow and scale. These hubs offer essential resources and guidance as companies navigate the complexities of the market.

6.4 VHT managers

6.4.1 Policy & Law makers

Policy and law makers are **governing and funding actors**, who draft and enact governance mechanisms through policies, legislations and public funding that regulate yet incentivize the development of VHT technologies for the safety, wellbeing and benefits of the citizens and the wider society. They are enablers that formulate policies, guidance and recommendations for fostering the entire VHT ecosystem that aligns all actors of the VHT CoP to work towards the common goal.

- **Healthcare policymakers** who may develop specific policies linked to the use of Digital Twins and may as well use this technology to support policy making, usually with the mediation of experts.

⁶³ <https://www.european-health-data-space.com/>

- Example: National ministries of health (NMH), EC institutions.
- **Healthcare lawmakers** at regional, national and European levels may establish a novel (or integrate the current) legislative framework governing the development and use of DT across the health sector, through enactment of necessary clauses, clarifications and provisions into existing regulations or those under revision.
 - Example: Members of European parliament, Members of regional and national governments.

6.4.2 *Regulators, standardization actors, testing and clinical trial actors*

Regulators, standardization actors, testing and clinical trial actors are **governing actors of VHT** who implement the governance instruments set out to test, standardize, regulate **and facilitate the market access of VHT technologies developed by the VHT developers**, such that the safety and efficacy of the technologies are assured and do not cause harm to users and beneficiaries (patients and citizens). They formulate guidelines to interpret and apply the legislative acts that govern VHT technologies.

Regulatory actors

- **Medical Device regulators** (e.g., FDA) that provide marketing authorization for Digital Twins that are used as clinical decision support systems, but also qualification for Digital Twins used as medical device development tools;
- **Drug/Advanced Therapy regulators** (e.g., FDA, EMA) that provide qualification for Digital Twins used as drug/ATMP development tools;
- **Conformity assessment agencies** (e.g., Notified bodies) designated by an EU country to assess the conformity against relevant regulation and applicable standards of Digital Twins for healthcare or of other medical products developed with Digital Twins before they are being placed on the market;
- **National competent authorities (NCA)** are responsible for authorisation of medicines that do not pass through the centralized procedure. NCA's can review or authorise the use of HDT to augment clinical trials, with virtual populations;
- **Competent authorities for medical devices (CAMD)** facilitates the implementation and enforcement of the Regulations on medical devices and on *in vitro* diagnostics, whereby they oversee the use of Digital Twins to demonstrate the safety, efficacy and effectiveness of medical devices, as well as drafting guidance on implementing the relevant regulations and standards;
- **Coordination groups**
 - Example: Medical Device Coordination Group (MDCG), Health Technology Assessment coordination group (HTA-CG).

Regulatory agencies play a crucial role in providing guidance, establishing standards, and promoting the safe and effective use of computational modelling and simulation in healthcare. The **US Food and Drug Administration (FDA)** has been actively involved in exploring the use of *in silico* methodologies and computational modelling in regulatory decision making. They have developed guidance documents and initiatives such as the FDA 2016 guideline on Reporting of Computational Modelling Studies in Medical Device Submissions⁶⁴, the FDA 2018 guidance on PBPK models⁶⁵, and the FDA draft guidance document⁶⁶ that outlines a generalised framework for assessing model credibility that relies heavily upon the ASME VV-40 standard⁶⁷. The **European Medicines Agency (EMA)** has also recognized the potential of *in silico* methodologies and computational modelling. They have established the "Innovation Task Force" to facilitate the uptake of innovative methods in the development of medicines. They published some guidance documents on the topic such as the EMA guideline on reporting PBPK models⁶⁸ and the EMA guidance for "Qualification of novel methodologies for medicine development"⁶⁹.

⁶⁴ <https://www.fda.gov/media/87586/download>

⁶⁵ <https://www.fda.gov/media/101469/download>

⁶⁶ <https://www.fda.gov/media/154985/download>

⁶⁷ <https://www.asme.org/codes-standards/find-codes-standards/v-v-40-assessing-credibility-computational-modelling-verification-validation-application-medical-devices>

⁶⁸ https://www.ema.europa.eu/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpb-modelling-simulation_en.pdf

⁶⁹ https://www.ema.europa.eu/documents/regulatory-procedural-guideline/qualification-novel-methodologies-drug-development-guidance-applicants_en.pdf

Different regulatory pathways and strategies for the acceptance of *in silico* methodologies and Digital Twins in medical device and drug development can be identified, including:

- **Certification of a SaMD:** Nowadays, regulatory authorities widely acknowledge software designed for medical purposes as a distinct category of medical devices known as Software as a Medical Device (SaMD) or Medical Device Software (MDSW). Both the FDA's Centre for Devices and Radiological Health (CDRH) and the European Union's CE-marking process have established regulatory pathways specifically designed for these types of technologies. Notably, there is a particular focus on SaMDs with predictive capabilities. An example class of SaMD includes the HeartFlow service solution, discussed in a previous chapter.
- **Qualification of *in silico* methodologies:** The FDA and EMA offer qualification pathways for medical device and drug development tools. While the FDA provides qualification for both, the EMA currently only provides it for drug development tools. The process for qualifying a new methodology is not mandatory but highly recommended; it involves requesting qualification advice from the regulatory authority followed by a formal request for qualification opinion. If a positive qualification opinion is obtained and there are no criticisms from the experts, the developer can utilise the methodology to generate evidence in a marketing authorization application for a new medical product.

Notified bodies are commercial organizations designated by an EU country to **assess the conformity** of certain products before being placed on the market. These bodies carry out tasks related to conformity assessment procedures set out in the applicable legislation, when a third party is required. They are assigned an important role in various health related regulations, such as the Medical Devices Regulation, the *In vitro* Diagnostic Medical Devices Regulation and the AI Act. A full list of notified bodies is available on the website of the European Commission⁷⁰. Umbrella organizations bringing together multiple notified bodies, such as **Team-NB**⁷¹, aim to streamline communication with the European Commission and assist their members in the implementation of new regulations and innovations.

Standardization actors

- **International standards bodies** are organizations that develop and publish technical standards intended for global use. These standards ensure quality, consistency, safety, and interoperability across different countries and industries.
 - Example: ISO, IEC, CENELEC
- **National standards bodies** are organizations responsible for developing, coordinating, and promoting standards within a specific country. These bodies ensure that national standards align with international standards to facilitate trade, enhance product quality, and ensure safety. They often represent their country in international standards organizations like the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC).
 - Example: DIN, ASME
- **International harmonization forums** are collaborative platforms where regulatory authorities, industry representatives, and other stakeholders come together to develop and align standards and guidelines across different countries and regions. These forums aim to ensure consistency, safety, and quality in various sectors by harmonizing technical requirements and regulatory practices across the globe.
 - Example: IMDRF
- **Community-driven standardization** are collaborative platforms where individuals, organizations, and communities come together to develop and promote standards based on shared needs and goals. Unlike traditional top-down approaches, these networks emphasize grassroots participation and collective decision-making. They often focus on creating community standards that reflect the specific requirements and values of the community involved.
 - Example: COMBINE-network

The standardisation actors are either grass-root standardisation initiatives or official committees of **standard defining organizations** (SDOs), which define technical or clinical standards. The most

⁷⁰ <https://webgate.ec.europa.eu/single-market-compliance-space/notified-bodies>

⁷¹ <https://www.team-nb.org/>

important international SDOs for technical standards are the **International Standards Organisation**⁷² (ISO) with its Technical Committees 215 (Health Informatics) and 276 (Biotechnology) and the **CEN / CENELEC**⁷³ with its technical committee 251 (Health Informatics). Several national SDOs are very active on the international stage, such as the American Society of Mechanical Engineers⁷⁴ (**ASME**) and the German standards organization **DIN**⁷⁵. The **Joint Initiative Council for Global Health Informatics Standardization**⁷⁶ (JIC) is an example of an umbrella organization formed to further the important role of health informatics standards to enable interoperability of information and processes across health domains and facilitate effective global markets for health information systems. In addition, there are several other institutions, societies and authorities defining standards relevant for VHTs. Beside that there are also some **grass-root communities** active in defining community standards, like COMBINE and the Genome Alliance for Genomics and Health⁷⁷ (GA4GH). A full list of SDOs defining technical standards is given in Table A.12 in Annex 2.

There are also several institutes, organizations, and initiatives which define clinical data standards. One of the most prominent of them is the Clinical Data Interchange Standards Consortium⁷⁸ (CDISC). A comprehensive list of SDOs relevant for defining clinical data standards is given in Table A.13 in Annex 2.

Testing and clinical trial actors

- **CROs** (Contract Research Organizations) that assist not only in the conduction of clinical studies for the validation of Digital Twins but also that can use Digital Twins to design and optimize clinical studies of new treatments;
- **Testing labs** (e.g. Reference laboratories)

6.4.3 Payers, Buyers, Reimbursement decision-makers, Health Technology Assessors

Payers, buyers, reimbursement decision-makers, as well as Health Technology Assessors are funders and governing actors of VHT who implement the governance instruments set out to evaluate the evidence to support the **clinical and economic effectiveness of VHT technologies** developed by the VHT developers, such that the safety and benefit of the technologies are assured brings value and benefits to healthcare systems and broader society (patients and citizens). They are enablers, who streamline the VHT ecosystem, such that the VHT creators/developers from the supply-side of the VHT ecosystem, meet the needs of the demand side from healthcare systems and society.

- **Payers** are typically seen as funders, public/private/statutory/national insurance bodies and healthcare authorities, who are collectively key **decision makers on coverage and reimbursement of Digital Twins in healthcare technology** that is offered by private or public healthcare providers / marketplace.
 - Example: **National competent authorities for pricing & reimbursement (NCAPR)** and public healthcare payers, statutory insurance bodies (public, private insurance);
 - Key players: European Social Insurance Platform (EUSIP), International Association of Mutual Benefit Societies (AIM), Association of mutual and cooperative insurers in Europe (AMICE), Federal Joint Committee (G-BA, Germany).
- **Buyers** are typically healthcare providers (hospital networks) and national healthcare systems that buy or procure Digital Twin healthcare technology to provide healthcare. This category represents a wide range of stakeholders linked to the national or regional healthcare provision model. It may include private and public **healthcare providers**, insurance, healthcare authorities, **group buyers**, etc.
 - Individual healthcare systems, healthcare providers, as well as the above payers representing the national competent authorities on pricing and reimbursement, are likely to perceive the role of buyers of Digital Twin technologies from clinician-facing DTs for clinical decision support to patient-facing DTs that support care pathways.
 - Group buyers are a collective of individual healthcare institutions or health systems or for example a group of university hospitals, who form consortium of buyers, who collectively identify a unmet clinical or healthcare system need, for which public procurement of innovations are explored. Example: European University Hospital Alliance (EUHA).

⁷² <https://www.iso.org>

⁷³ <https://www.cenelec.eu/>

⁷⁴ <https://www.asme.org/>

⁷⁵ <https://www.din.de>

⁷⁶ <http://www.jointinitiativecouncil.org/>

⁷⁷ <https://www.ga4gh.org/>

⁷⁸ <https://www.cdisc.org/>

- **Decision makers** are a collective of payers and buyers at national level of member states, who commission and evaluate the assessments made by the HTA bodies (see below), to decide upon the recommendations on the payments / coverage of Digital Twins destined as individual DT technologies or when they are part of a care pathway. National decision-making agencies examine aspects like risk-benefit, clinical-effectiveness and cost-benefit of Digital Twins, while taking into account the value it brings for their population and healthcare system.
 - The recommendations of the decision maker group on the effectiveness and value-benefit of Digital Twin solutions, form the basis of negotiation between the developers and funders/payers/buyers of any health technology.
 - Example: in Germany the federal council GBA, calls for the technology assessment through IQWiG, the German HTA body, whose report helps define the recommendation on the coverage decisions.
- **Health technology assessment bodies** are actors that assess the clinical effectiveness, cost-effectiveness, and reimbursability level of new Digital Twins for healthcare or of other medical products developed with Digital Twins. This group includes Health technology assessors, Health economists and HTA committees that include representatives of medical societies and patient organizations.
 - On the request of payers/buyers or technology developers, they are often tasked to evaluate the evidence on value of Digital Twin solutions, for an unmet healthcare need or against the standard of care setting within national healthcare systems;
 - Example: Institute for Quality and Efficiency in Healthcare (IQWiG, Germany), National Institute for Health and Care Excellence (NICE UK).
- **Healthcare innovation procurement experts** are the specialised actors that empower the key end-users (patients, clinicians), the buyers (healthcare network, hospitals) and the payers (healthcare systems, insurance forum), to proactively explore the strategic procurement of healthcare innovations like VHT technologies that solves a current healthcare need, for which a solution is not current available in the off-the shelf marketplace.
 - In the context of the VHT ecosystem, innovation procurement experts could act as key liaisons who can bring the incentive to the VHT ecosystem, as they look for a shared-risk model with DT developers. For instance, through pre-commercial procurement, they facilitate the procurement of research that leads to joint development and testing the low-TRL (Technology Readiness Level) DT solutions that cater to their unmet need. Thereby channelize the VHT ecosystem to develop innovations that are better and faster prepared for regulatory and reimbursement assessment.
- **Independent research institutes, public institutions (regional and national), hospital-based HTAs** that are likely to conduct evidence-based reviews of Digital Twins for benefit-risk assessment and to measure their value and pricing.
 - Examples: Institute for clinical and economic review (ICER⁷⁹), Sciensano⁸⁰, Organisation for Economic Co-operation and Development (OECD⁸¹), European Health Observatory on Health Systems and Policies⁸² – a partnership hosted by World Health Organisation (WHO) that identifies and generates the evidence on health systems for Europe’s decision-makers, WHO-Europe⁸³.
 - Initiatives: EDiHTA⁸⁴, ASSESS DHT⁸⁵.

With respect to the reimbursement for DTs in healthcare, the roles and the respective coverage decision pathways are set to vary for VHT technologies, subject to categorization of the DT as health technology within the spectrum of drugs, medical devices, drug-device combination or digital health technologies.

⁷⁹ Institute for clinical and economic review - ICER - pioneering leaders to conduct comparative effectiveness review in U.S., whose reports are most sort after - <https://icer.org/who-we-are/>

⁸⁰ Sciensano – Health service Research Organization - Belgium, who conduct health services research, to examine healthcare system to improve health and quality of life of patients - <https://www.sciensano.be/en/about-sciensano/sciensanos-organogram/health-services-research>

⁸¹ OECD - Organisation for Economic Co-operation and Development, policy think tank who conduct independent research on measuring health outcomes and health system resources, as well support evidence-based policies that improve access, efficiency , and quality of healthcare. - <https://www.oecd.org/en/topics/health.html>

⁸² European Health Observatory on Health Systems and Policies – a partnership hosted by World Health Organisation (WHO) the identifies and generates the evidence on health systems, for Europe’s decision-makers - <https://eurohealthobservatory.who.int/>

⁸³ WHO-Europe - Responsible Authority for public Health with United Nations system, supports countries in meeting citizen’s expectation about health - <https://www.who.int/europe/about-us/about-who-europe>

⁸⁴ EDiHTA – The first European Digital Health Technology Assessment framework - <https://edihta-project.eu/>

⁸⁵ ASSESS DHT – Development and harmonization of methodologies for assessing digital health technologies in Europe - <https://assess-dht.eu/>

Several HTA agencies have demonstrated interest in *in silico* methods and have proposed related initiatives. Here are some representative examples:

- National Institute for Health and Care Excellence⁸⁶ (NICE) - United Kingdom: NICE has expressed interest in *in silico* methods and their potential impact on HTA. They have highlighted the use of computer modelling and simulation techniques to support decision making, particularly in areas such as medical devices and diagnostics. They presented a 5-year strategic plan where the use of digital health technologies is among the 6 key identified trends⁸⁷.
- French National Authority for Health⁸⁸ (HAS) - France: HAS has been exploring ways to integrate computer modelling and simulation techniques into their assessment processes and has developed a classification system for digital solutions used in healthcare based on three criteria: intended use, capacity to provide a personalised response, and autonomy in decision-making⁸⁹.
- AQuAS⁹⁰ (Agency for Health Quality and Assessment of Catalonia) is an organization based in Catalonia, Spain, dedicated to assessing the quality and effectiveness of healthcare interventions and technologies. AQuAS conducts HTA studies to inform decision-making processes related to healthcare policy, resource allocation, and reimbursement decisions. The agency plays a significant role in evaluating the clinical and economic aspects of healthcare technologies, including pharmaceuticals and medical devices. AQuAS recently contributed to the discussion on the role of AI within *in silico* medicine⁹¹.

At the international level, one of the most important organizations dedicated to health technology assessment is the **Health Technology Assessment International (HTAi)**⁹². HTAi brings together experts, professionals, and organizations involved in HTA from around the world. It serves as a global platform for knowledge exchange, collaboration, and development of HTA methodologies and practices. HTAi organises annual conferences and provides opportunities for networking and collaboration among HTA stakeholders. They have organised workshops and sessions to discuss the role of *in silico* trials to complement or supplement randomised controlled trials of new technologies. At the EU level, **EUnetHTA**⁹³ was established to create an effective, collaborative and sustainable network for HTA across Europe – to help develop reliable, timely, transparent, and transferable information to contribute to HTA in European countries.

6.4.4 Legal, IP, Ethical, & Social actors

Legal, IP, Ethical and Social (ELSI) actors include national and EU Parliaments, Competent supervisory Authorities and Courts, Member States and European regulatory and harmonization bodies **guiding and regulating the VHT ecosystem** by addressing the legal, ethical, social, and intellectual property aspects of the VHT. These actors lay down the applicable legal framework and ensure compliance with laws, foster ethical decision making, and address societal implications to build trust and acceptance of VHT solutions. They also include legal counsels, Data Protection Officers, compliance officers, AI and social science experts and ethicists, supporting the VHT community in navigating complex legal and social frameworks, while promoting equity, inclusiveness, and moral responsibility in the development, deployment, and application of Digital Twin technologies.

- **GDPR officers**, including data controllers, data providers, data protection officers (DPO), supervisory authorities, and their legal advisors, involved with the handling of sensitive data used to develop, validate, or use Digital Twins.
- **Ethics experts, members of ethical committees in research, clinical, policy or public domain** would evaluate the benefits and risks introduced by Digital Twin technologies. Combined with the legal pre-requisites enshrined in law, ethical obligations identified by these experts would foster moral considerations that shape the action of the VHT community, impose limits and offer guidance (*e.g.* VHT Code-of-conduct) to embrace and navigate Digital Twin technologies.
- **Social science experts specialised in Science & Technology studies**, would evaluate the relationship between the different actors within the VHT community of practice, to unearth the frictions on social acceptance of Digital Twins in healthcare. They offer insights and guidance to promote acceptability and

⁸⁶ <https://www.nice.org.uk/>

⁸⁷ <https://static.nice.org.uk/NICE%20strategy%202021%20to%202026%20-%20Dynamic,%20Collaborative,%20Excellent.pdf>

⁸⁸ <https://www.has-sante.fr/>

⁸⁹ https://www.has-sante.fr/upload/docs/application/pdf/2019-07/rapport_analyse_prospective_20191.pdf

⁹⁰ <https://aquas.gencat.cat/ca/inici>

⁹¹ Geris L *et al.* VPH Institute & Avicenna Alliance White Paper (2022). <https://www.doi.org/10.5281/zenodo.8064147>

⁹² <https://htai.org/>

⁹³ <https://www.eunethta.eu/>

trust in the developed technologies, in order to increase collaboration towards socially responsible Digital Twin technology solutions.

6.5 VHT catalysts

6.5.1 Professional associations, Trade bodies, Civil society organizations

Professional associations serve as crucial boundary spanners within the VHT ecosystem, facilitating cross-fertilization among diverse stakeholders. By **fostering a shared understanding of common goals and priorities**, these associations play a vital role in guiding the community towards collective success. By acting as local champions, professional associations inspire their members to embrace the VHT revolution. They effectively bridge the gap between the community and broader stakeholders by gathering and disseminating valuable insights, including emerging challenges and unanswered questions.

- **Professional associations such as scientific/clinical societies** already play multifaceted roles within the healthcare landscape. Their members often go beyond their primary responsibilities (*e.g.*, patient care) to serve on Ethics Research Committees, Health Technology Assessment committees, and regulatory assessment panels. To effectively facilitate the integration of Digital Twins into various professional domains, these associations will play a formidable role in educating and empowering their members to embrace and effectively utilize this transformative technology.
 - Example: European Society for Cardiology (ESC).
- **Trade bodies** play a crucial role within the CoP by providing essential industry insights and driving the industrialization of Digital Twin technologies. As representatives of thousands of SMEs, they bring valuable experience in adopting and implementing new technologies within clinical practice. Furthermore, trade bodies tend to actively seek partnerships with other community subgroups to gain a deeper understanding of the ethical, legal, and social implications of VHT innovations, while offering their support and expertise on technology innovations.
 - Example: MedTech Europe, European Federation from Pharmaceutical Industries and Associations (EFPIA)
- **Expert groups, Think Tanks, Policy Forums, Independent research institutes and Networks** play a critical brokering role within and outside the VHT ecosystem. Leveraging their independent stature, these organizations conduct assessments of healthcare systems and advise governments on policies and implementation strategies to integrate VHT solutions effectively. These organizations are particularly crucial for widening the VHT technologies to developing countries, which rely on in-depth assessments from these pioneering bodies to efficiently allocate resources and maximize the impact of investments in VHT solutions on the well-being of their populations.
 - Example: National Academies of Science Engineering & Medicine (USA), OECD, WHO, European Health Observatories, EIT-Health⁹⁴, European Health Telematics Association (EHTel⁹⁵)
- **Civil Society organizations** are crucial stakeholders in the VHT ecosystem, representing the voice of the public and holding policymakers accountable. As key drivers of societal change, they are instrumental in shaping the future of healthcare. Gaining the trust and active engagement of CSOs is paramount for the successful implementation of DT technologies. The VHT CoP should actively support CSOs by providing them with comprehensive information about the potential benefits and risks of VHT solutions. Furthermore, the VHT CoP should actively seek input from CSOs, recognizing their crucial role in shaping the ethical and societal implications of these emerging technologies. By actively involving CSOs in the development and implementation of VHT solutions, the community can ensure that these technologies are developed and deployed in a responsible and equitable manner.
 - Example: European Patients' Forum, EU4Health

6.5.2 Education, training, communication actors

Education, training, communication actors **create educational resources**, develop and disseminate high-quality educational materials, including lectures, online courses, and interactive tools, to broaden public understanding and facilitate self-learning about Digital Twin technologies. One important educational goal is to **develop the future workforce**: train the next generation of researchers, developers, and practitioners in the creation, utilization, and interpretation of Digital Twin technologies.

⁹⁴ <https://eithealth.eu/>

⁹⁵ <https://www.ehtel.eu/>

Another important goal is to **educate the public**, to inform and engage the broader public about the potential benefits and implications of Digital Twin technologies, fostering understanding and acceptance within society.

- **Teaching staff** who use Digital Twins during their lectures for education and demonstration.
- **Teachers and Trainers** that offer lectures in universities and summer schools, to train the next generation on developing, applying or using the Digital Twins.
- **Accredited professional training institutes** that (re-)train multidisciplinary professionals in the clinical, industrial or regulatory sectors, about the interplay of their professions and the lifecycle of VHT technologies. For example, training and retraining programs for professionals in clinical, regulatory and industrial sectors, to interpret VHT-based results, is seen as paramount for capacity building and sustainability of the VHT ecosystem.

Box 11: Success story –the Immune Digital Twin initiative

Success story: ecosystem building

Website: <https://www.rd-alliance.org/groups/building-immune-digital-twins-wg>

The **Immune Digital Twin (IDT) Project** is an interdisciplinary initiative focused on developing personalized computational models of the human immune system to enhance precision medicine. Recognizing the immune system's complexity and its pivotal role in various diseases, the project aims to create digital replicas that can predict individual immune responses, thereby improving diagnostics and treatment strategies.

The initiative started by a 3-week workshop at the Institut Pascal⁹⁶ bringing together close to 100 experts in the field. The core team continued to build the community and the IDT has now been accepted and endorsed as a Research Data Alliance working group. This will provide the community with a limited level of support (logistical, not financial) to realize its proposed activities. The WG's primary objectives include:

- **Literature Repository:** Developing open-access data and model repositories to serve as shared resources for the IDT community. This involves creating metadata libraries that curate relevant literature and existing models of the human immune system.
- **Data Infrastructure:** Establishing robust, shareable, and scalable data infrastructures that adhere to FAIR (Findable, Accessible, Interoperable, Reusable) principles. This infrastructure will support the integration and expansion of computational models into comprehensive medical Digital Twins.
- **Community Building:** Fostering a long-term interdisciplinary community dedicated to addressing the challenges inherent in developing IDTs. This includes organizing webinars, workshops, and collaborative projects to facilitate knowledge exchange and innovation.

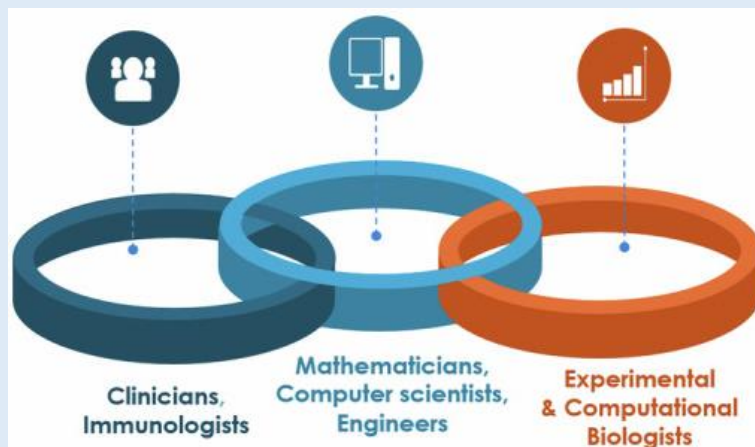


Figure 4: Interdisciplinarity required for building immune Digital Twins⁹⁷

While computational modelling is a critical component, the development of medical Digital Twins extends beyond modelling to include comprehensive data integration and standardization. By focusing on the immune system, the project aims to create a foundational framework that can be adapted to other biological systems in the future. Through these efforts, the IDT Project aspires to revolutionize personalized medicine by providing clinicians with advanced tools to simulate and predict patient-specific immune responses, ultimately leading to more effective and tailored healthcare solutions.

⁹⁶ <https://indico.ijclab.in2p3.fr/event/9017/>

⁹⁷ <https://doi.org/10.1038/s41540-024-00450-5>

7 Needs, Challenges and Barriers to the Virtual Human Twin

7.1 Analysis of stakeholder needs and challenges

In a wide range of stakeholder engagement activities, the needs and challenges of stakeholder groups' involvement with the VHT were identified. A detailed SWOT analysis for different stakeholder groups, summarized below, assessed the strengths, weaknesses, opportunities and threats with respect to the VHT development and deployment for that particular stakeholder group.

7.1.1 Patients & Citizens

Patients and citizens acknowledge the importance of **data sharing as a cornerstone of the VHT** paradigm but emphasize the need for assurances regarding the ethical and responsible use of their health data. While there is a **willingness to contribute** their personal information, they expect robust data governance frameworks that prioritize privacy and security while enabling secure and efficient data utilization for the advancement of VHT technologies. Moreover, concerns regarding artificial intelligence (AI) and synthetic data used in VHT development remain prevalent. Specifically, patients and citizens are apprehensive about biases inherent in AI algorithms and have questions regarding the reliability and validation processes of synthetic data. To address these concerns, transparency and explainability in AI systems are essential, alongside clear communication about the role of synthetic data in supplementing real-world datasets.

Another pressing issue is the need to **enhance patient and citizen understanding** of VHT and digital health technologies. Many individuals feel unprepared to fully comprehend the implications of VHT for their healthcare and call for accessible educational resources that improve digital health literacy. Such initiatives should **empower individuals** to participate actively in the VHT ecosystem and make informed decisions regarding their healthcare. Furthermore, equitable access to VHT technologies is a significant concern. Patients and citizens highlight the **risk of disparities** in access based on socioeconomic status and geographic location. Efforts must be directed toward promoting the development and implementation of VHT solutions that are universally accessible and address potential biases in algorithms and datasets to avoid perpetuating healthcare inequities.

The SWOT analysis underscores the potential of VHT to empower patients, enhance personalized care, and accelerate medical research as key strengths. Weaknesses include limited public awareness, risks to privacy, and the possibility of exacerbating existing health disparities. Opportunities are identified in fostering patient-centred design, using VHT for education and empowerment, and establishing robust ethical frameworks to guide implementation. Threats such as data misuse, distrust in AI technologies, and inadequate regulatory oversight remain challenges to be addressed.

7.1.2 Clinical Community, Healthcare Providers, and Hospitals

The clinical community and healthcare providers **recognize the transformative potential of VHT** but stress the **importance of demonstrating its clinical relevance** and value in real-world healthcare settings. Clinicians require clear evidence of VHT's impact on patient outcomes, cost-effectiveness, and workflow efficiency to support its integration into clinical practice. Rigorous clinical trials and studies tailored to specific healthcare applications are essential to validate the utility of VHT. In addition, clinicians often face challenges in interpreting the complex data generated by multi-scale Digital Twins. This necessitates the development of user-friendly interfaces and visualization tools that simplify data presentation and ensure that healthcare professionals can effectively engage with the technology. Adequate **training programs and support systems** are also required to equip clinicians with the skills needed to utilize VHT in their practices.

Integration of VHT into existing clinical workflows presents technical and logistical challenges, particularly in ensuring interoperability with electronic health records (EHRs) and other healthcare information systems. **Collaboration** with healthcare providers to design solutions that align with existing workflows is critical to seamless implementation. Data quality, security, and patient privacy also emerge as key concerns in VHT deployment. Robust **data governance frameworks** and privacy-preserving techniques are needed to protect patient data and maintain confidentiality while addressing biases in algorithms and datasets to ensure fairness and equity.

The SWOT analysis highlights the strengths of VHT in improving diagnostics, enabling personalized treatment plans, and enhancing clinical decision-making. Weaknesses include limited availability of tailored VHT solutions, a lack of an extensive evidence base to support widespread adoption, and the potential for technology-induced errors. Opportunities lie in fostering partnerships between clinicians and developers, integrating VHT into medical education, implementing adequate change management processes and addressing unmet clinical needs. However, threats such as apprehension to adopting new technologies, interoperability challenges, and ethical concerns regarding patient autonomy and data privacy must be mitigated.

7.1.3 Academia

In academia, researchers face significant obstacles in securing **access to diverse, high-quality datasets** that are representative of various patient populations and disease states. These datasets are critical for the development and validation of VHT models. To address this issue, mechanisms for data sharing and collaboration among research institutions, healthcare providers, and industry partners must be established. Another challenge is the **findability, accessibility, and interoperability of existing VHT tools and resources**. The lack of standardization and centralized repositories impedes researchers' ability to locate relevant software, models, and datasets. Developing standardized frameworks for metadata, ontologies, and common data formats is essential to enhance the usability of VHT resources. The **computational demands** of VHT development represent another significant barrier. Developing and simulating complex VHT models requires substantial computational resources and specialized expertise, which are often inaccessible to many researchers. Continued investments in high-performance computing infrastructure and accessible cloud-based platforms are necessary to support these efforts. Collaboration across disciplines is also a priority, given the inherently multidisciplinary nature of VHT research. Platforms and forums for **knowledge sharing, training, and interdisciplinary engagement** should be established to foster growth within the VHT research community. Additionally, there is a need to **incentivize** academics to engage with and contribute to the VHT infrastructure. Clear rewards and recognition systems, such as citation metrics, peer review acknowledgment, and funding opportunities, must be implemented to motivate researchers to dedicate time and resources to this field. The SWOT analysis identifies challenges such as limited data availability, difficulties in locating and integrating tools, computational resource constraints, and the need for interdisciplinary collaboration. However, opportunities exist in fostering collaboration, improving data access, and incentivizing academic contributions. Overcoming these barriers will ensure that academia plays a pivotal role in advancing VHT technologies.

7.1.4 Industry

The industry sector faces uncertainty in **navigating the evolving legal and regulatory landscape** for VHT-based products and services. The lack of clear guidelines hinders commercialization and creates uncertainty for companies. To address this, advocacy for streamlined regulatory pathways, standards, and legal frameworks is essential to influence the development of commercial VHT applications. **Demonstrating the value proposition** of VHT to healthcare providers and payers is another key challenge. Companies must provide compelling evidence of clinical and economic benefits through robust clinical trials and economic evaluations. Effective communication materials are critical to conveying these benefits to healthcare decision-makers.

Data security, privacy, and ethical considerations remain central to the successful deployment of VHT solutions. Questions surrounding liability and intellectual property when using the VHT infrastructure need to be addressed. Companies must implement robust data governance frameworks and address biases in algorithms and datasets to ensure ethical and equitable applications. The **shortage of skilled professionals** with expertise in VHT technologies presents an additional barrier. Investments in training programs and partnerships with academic institutions are necessary to develop a skilled workforce capable of supporting VHT development and implementation.

The SWOT analysis reveals that VHT offers significant strengths, including potential cost savings, faster product development, and personalized healthcare solutions. However, high development costs, scalability issues, and lack of reimbursement models are notable weaknesses. Opportunities lie in leveraging VHT for drug discovery, clinical trial optimization, and personalized medicine, while threats include regulatory uncertainty, data privacy concerns, and workforce shortages.

7.1.5 *Regulatory, Standardization, Legal, Ethical and Social Sciences Actors*

Regulatory and standardization actors face the challenge of establishing clear and consistent guidelines for **data anonymization and pseudonymization**, particularly in light of varying interpretations of data protection regulations such as the GDPR. These inconsistencies hinder **cross-border data sharing** and VHT development. Harmonizing data protection regulations and developing clear guidelines are critical to overcoming this challenge. The **regulatory classification of VHT software components** remains ambiguous, creating uncertainty for developers and complicating approval processes. Establishing precise definitions and classifications is necessary to streamline regulatory pathways.

Standardization is another priority, as the absence of uniform data formats, ontologies, and workflows hampers interoperability and reproducibility. Collaborative efforts among regulatory agencies, standardization organizations, as well as academic and industry stakeholders are required to develop and adopt standards that align with regulatory requirements and developer needs. **Ethical concerns**, including algorithmic bias, patient autonomy, and the responsible use of patient data, must be addressed through robust ethical guidelines and frameworks. Establishing **social acceptance** and building trust in VHT technologies requires co-creation of these technologies with end users, following responsible research and innovation approaches.

The SWOT analysis emphasizes strengths such as the potential for harmonizing regulations and standards, while weaknesses include jurisdictional variations and a lack of established frameworks. Opportunities lie in promoting global collaboration and ethical standardization, but threats such as misaligned regulations and ethical disputes require careful navigation.

7.1.6 *Payer, Buyer, and Health Technology Assessment Community*

The payer and HTA community must establish clear criteria for **evaluating the clinical and economic value** of VHT solutions. Standardized methodologies are essential to assess clinical effectiveness, cost-effectiveness, and budget impact. Uncertainty surrounding the **long-term benefits and risks of VHT** further complicates coverage decisions. Research into the long-term effectiveness, safety, and cost-effectiveness of VHT is required to address this knowledge gap. Traditional **reimbursement** models may not suit the unique nature of VHT technologies, necessitating the development of innovative payment mechanisms, such as value-based payments and risk-sharing agreements.

Ensuring equitable access to VHT technologies is another priority. High costs could exacerbate healthcare disparities, necessitating policies that promote affordability and accessibility in underserved communities. Effective communication and **collaboration** between VHT developers, payers, and HTA bodies are crucial to align expectations and facilitate evidence generation.

The SWOT analysis highlights the payer community's experience in evaluating new health technologies as a strength, while identifying weaknesses such as a lack of clear guidelines, uncertainty about long-term benefits, and the absence of established reimbursement models. Opportunities include partnerships to generate evidence of VHT's value, innovative payment models, and improved population health outcomes. However, threats such as resistance to new technologies, potential healthcare cost increases, and ethical concerns regarding access and privacy must be addressed.

7.2 *Barriers to the adoption of the VHT*

Despite the consensus on the transformative potential of Digital Twins in Healthcare in medical research, healthcare delivery, and patient care, there are still substantial barriers hindering its widespread translation and adoption as discussed above. These barriers can be grouped into broader categories related to data, technology, ethical, legal and social consideration, economic and sustainability factors, as well as clinical uptake and adoption.

1. Data-related barriers

- **Data availability and quality:** the development and validation of reliable VHT models require large volumes of accessible, high-quality and diverse data. The lack of high-quality and properly annotated data poses a significant challenge.
- **Data collection and harmonisation:** there is a lack of standardized or even consensus-driven protocols for data collection and processing, hindering consistency and reliability in data production. Implementation of rigorous quality control measures at each stage of the data

lifecycle, coupled with the adoption of clear metadata standards, is required to further enhance data transparency and reproducibility.

- **Data interoperability and standardization:** integrating data from disparate sources with varying formats and standards is crucial for building comprehensive VHTs. Lack of interoperability and standardization in these collaborative research environments where data is integrated, can lead to potential errors, misinterpretations, and compromised outcomes.

2. Software and Infrastructure Barriers

- **Software tools and models:** the VHT infrastructure needs robust software tools and models for DT development, simulation, visualization, and data management. These tools and models should follow the FAIR (findable, accessible, inter-operable and reproducible) principles to avoid expending valuable resources reinventing the wheel.
- **Model complexity and integration:** developing and validating multi-scale and/or multi-organ VHT models that accurately represent the complexity of human physiology and pathology can be computationally intensive and requires advanced modelling and inference techniques. While running these integrated models through orchestrations of sub-models or data flows offers potential solutions, they face challenges in computational efficiency, interoperability, and tool architecture. In addition, there is a lack of widely used data and API standards for orchestrating the linking of multiple DTs.
- **Model complexity and run time:** the need to produce results within clinically realistic timescales to aid clinical decision making, provides a further challenge for integrated DTs, which can be addressed by creating surrogate models using AI.
- **Infrastructure:** the VHT infrastructure should be user-friendly and capable of handling large-scale data and complex models, while facilitating the interoperability with commercial platforms used in industry and clinics. The infrastructure should assist users in meeting the requirements regarding standardization and interoperability of their resources.

3. Ethical, Legal and Social Barriers

- **Data privacy and security:** the sensitive nature of health data necessitates strict privacy and security measures to ensure responsible and ethical data sharing. Widespread uncertainty regarding the concept of anonymous data and national divergences regarding legal bases on data reuse further complicate the scenario.
- **Informed consent and transparency:** clear guidelines and procedures for obtaining informed consent from individuals for various forms of data use are (remain) essential. Additionally, transparency regarding the underlying assumptions, algorithms and potential biases in VHT models is crucial for building trust with society at large.
- **Equity and access:** Ensuring equitable access to VHT technologies across diverse populations is crucial to avoid exacerbating health disparities.
- **Legal clarity and liability:** addressing legal clarity, certainty and liability is required to provide reassurance to developers, investors and healthcare providers regarding legal compliance, support and the allocation of responsibility in the event of any unintended consequences or harm arising from the use of this technology, including hospital down time due to *e.g.* security issues.

4. Regulatory Science Barriers

- **Lack of consolidated regulatory pathways:** the fragmentation of the regulatory decision-making landscape, and the absence of established standards and precedents for evaluating the safety and efficacy of the novel and complex VHT technologies, slows down their translation.
- **Lack of validation data:** the lack of sufficient, accessible and high-quality data for VHT development limits the ability to develop comprehensive and credible models that can be effectively used for research, regulatory assessment, clinical applications and commercialization.

5. Economic and Sustainability Barriers

- **Funding and investment:** sustained funding and investment are essential for advancing VHT research, developing and maintaining its infrastructure as well as supporting the innovation framework, providing innovation hubs and training VHT innovation officers.
- **Business models and value proposition:** clear business models and a compelling value proposition are needed to incentivize participation of all stakeholders along the VHT value chain (including patients) and attract private investments, ensuring the long-term sustainability of the VHT ecosystem.
- **Open source conundrum:** while open-source licensing could enhance collaboration and streamline development, competitive funding strategies, institutional plans for commercialization, and investor preferences for exclusive rights can hinder its adoption. Achieving a balance between fostering open-source solutions and maintaining viable business models, including mechanisms for developer participation in profits, is essential to ensure development and maintenance of a high-quality, sustainable VHT infrastructure.
- **Skilled workforce:** training and retraining of the workforce, from technology developers, over regulatory experts to healthcare providers, is required to ensure presence of the necessary competencies to develop, handle or assess VHT technologies.
- **Lack of specific reimbursement pathways:** the absence of established reimbursement pathways creates uncertainty, limiting the incentives for adoption. This challenge also affects the developers' ability to demonstrate financial value in the healthcare ecosystem.

6. Clinical Integration and Adoption Barriers

- **Clinical validation and utility:** demonstrating the clinical validity and utility of VHT applications through rigorous studies and trials, communicated in clinical publications and guidelines, is crucial to show sufficient safety and benefits for gaining clinical and hospital buy-in and integrating these technologies into clinical practice.
- **Integration into existing workflows:** the practical challenge of integrating these technologies into current clinical workflows, as well as provisioning access to already existing data generating tools (wearables, devices), often deters adoption.
- **Interdisciplinary teams:** teams of healthcare professionals and medical technology experts are needed to integrate and adopt the variety of devices, instruments or services that use or deliver VHT applications acquired by healthcare providers.

7. Cultural and Organisational Barriers

- **Resistance to change:** some stakeholders may resist adopting new technologies as they involve changes to already established workflows, or because of scepticism, lack of knowledge or mistrust. Implementing proper change management processes will allow one to prepare and support individuals, teams and leaders in making organizational changes, in order to embrace the VHT.
- **Education and training:** educating stakeholders, including patients, healthcare professionals, regulators and others, about VHT technologies, their applications and appropriate use, as well as their limitations is essential for fostering adoption and effective use in healthcare.
- **Geographical and organisational differences:** the use of VHT applications in the “real world” requires working with local constraints, which can differ among countries, among regions in a single country or even among organisations within the same region.

Addressing these barriers requires a multi-faceted approach involving collaboration among researchers, clinicians, patients, industry, policymakers and regulatory bodies. Overcoming these challenges is essential towards unlocking the transformative potential of VHTs and realizing a future of personalized, predictive, and preventative healthcare.

8 Need and Vision for VHT: Conclusions and Recommendations

8.1 Conclusions

Taking into account the current healthcare, policy and innovation context, as well as the needs and challenges identified by the stakeholders across the VHT ecosystem, this first part of the roadmap proposed a high-level vision for the Virtual Human Twin.

A **Virtual Human Twin** is an integrated multi-level, -time, and -discipline digital representation of a body, organ, or cell enabling the comprehensive characterisation of the physiological and the pathological state in its heterogeneity, allowing patient-specific predictions for the prevention, prediction, screening, diagnosis and treatment of a disease, as well as the evaluation, optimisation, selection, personalization, and de-risking of intervention options.

The Virtual Human Twin initiative represents a groundbreaking approach to addressing key challenges in healthcare, including rising costs, the growing demand for personalized medicine, and the need for more efficient healthcare systems. By leveraging advanced *in silico* technologies such as computer modelling, simulation, and artificial intelligence, supported by state-of-the-art computational infrastructure and data technologies, the VHT enables the creation of virtual representations of human physiology and pathology. These virtual models offer insights into disease progression across scales of time, space, and organ systems, paving the way for personalized, predictive, and preventive care, accelerating medical research, and improving support for clinicians, patients, and healthcare providers.

The **VHT initiative** emphasizes both generating new scientific knowledge and applying it to enhance clinical care. Its realization requires collaborative efforts at a European level, including the establishment of a dedicated VHT infrastructure. In order to achieve its core goals, the VHT initiative furthermore requires work on pooling resources and expertise, data sharing and interoperability, credibility and trust, as well as economic sustainability. As such the VHT initiative will provide a systematic, ever-growing digital and quantitative representation of the actionable knowledge available on human pathophysiology. Its federated public infrastructure will enable the pooling of resources and assets (data, models, algorithms, computing power, storage etc.) to develop Digital Twins in healthcare and assess their credibility. A **collaborative, diverse, and engaged ecosystem** will facilitate the uptake of the VHT across all sectors and R&D phases, including clinical practice, with full respect of ethical, legal, and social considerations.

Realizing the vision of the VHT, requires development on a wide range of topics, including technology, infrastructure, standards, regulatory processes, ELSI (ethical, legal and social issues), users, uptake, as well as business models and sustainability. Collaboration between researchers, clinicians, policymakers, and industry stakeholders is essential to establish robust technical solutions, ethical guidelines, and regulatory frameworks.

8.2 Recommendations

The essential starting point for realising this vision is engaging the innovation ecosystem, identifying all relevant stakeholders and assessing their expectations, needs and challenges. The information presented in this part can be summarized in the following recommendations for the VHT.

1. **Research and innovation:** to ensure that the Virtual Human Twin reaches its full potential, research and innovation (R&I) are paramount. This R&I should span a continuum from basic research, addressing fundamental knowledge gaps and the development of nascent technologies, to translational research, bridging the gap between laboratory findings and clinical applications. The scope of R&I should cover the creation of generic and population-specific Digital Twins as well as highly personalized Digital Twins, catering to diverse applications. Embracing blue-sky research will be crucial for breakthroughs and innovation, alongside the exploration of diverse use cases to demonstrate the VHT's broad applicability.

2. **Prioritizing use cases to enhance VHT impact:** the identification, development, and delivery of high-impact validated use cases is important to demonstrate the practical value of the VHT across diverse clinical and scientific applications. Prioritization should be given to use cases that address key areas such as diagnostics, medical education and training, clinical decision support, therapy development, and intervention planning. The development of a demonstrator (such as the Virtual Brain Twin) showcasing the capabilities of the VHT in a specific area will be valuable for engaging stakeholders and showcasing its potential. Selecting a diverse range of use cases will illustrate the VHT's versatility and broad applicability, particularly highlighting the transition from screening diseases to more complex stratified simulations. This approach ensures that the VHT is developed with a focus on real-world applications, maximizing its impact and adoption.
3. **Enhancing clinical usefulness - a cornerstone for VHT adoption:** to ensure the VHT is embraced in clinical practice, demonstrating its clinical usefulness is paramount. This requires a rigorous assessment of its clinical effectiveness and usability. Following a co-creation approach, the conceptualisation and development of the VHT must be driven by a clear understanding of clinical needs, ensuring the VHT addresses practical challenges and provides tangible benefits. Fostering dialogue and collaboration with clinicians throughout the development process will ensure the VHT is aligned with clinical workflows and priorities.
4. **Applications across the disease continuum - realizing the full potential of the VHT:** to maximize the impact of the Virtual Human Twin, it is essential to advance the understanding of how its solutions, products, and services can be applied across the entire disease continuum. This includes exploring its potential in prevention, treatment, and follow-up, spanning a wide range of applications from biomedical and clinical studies to therapy development and diagnostics. The VHT should not be confined to traditional healthcare settings but also encompass innovative care models such as remote care and self-care, empowering individuals to actively manage their own health. This comprehensive approach will unlock the full potential of the VHT, transforming healthcare from a reactive system to a proactive, personalized, and patient-centric model.

PART 2:

REALISING THE VHT - TECHNOLOGY

9 Organisation of resources for VHT

9.1 Multidimensional space as an organisational paradigm

The elements discussed in Part 1 of this roadmap provide a general outline for the blueprint of the Virtual Human Twin. This section provides a high-level description of the essential characteristics of the VHT. From a technology perspective, the VHT can be thought of as a **collection of resources** that can seamlessly be interrogated, integrated, and executed, as well as the underlying fabric to facilitate such actions. Every resource that forms the VHT is extensively annotated, and this metadata can be used for advanced searches and other complex operations. However, it is crucial to identify an **organisational paradigm**, referring to a conceptual framework or overarching system of organization that defines how elements within a complex structure are categorized, related, and presented to users. The VHT's organisational paradigm is a guiding principle for structuring its resources, enabling intuitive navigation, effective data representation, and enhanced interaction within a multidimensional space⁹⁸.

The basic **VHT resources** are data and models, which can be processed by tools and integrated into scientific workflows. These resources are stored and executed in the VHT infrastructure made up of storage, execution, and network resources, organised and integrated into a repository/catalogue and simulation platform, relying on tailored software stacks.

The central resource in the VHT's organisation is the **data object**. Annotation is the labelling of this data with necessary information to understand, interpret, and use the data object. A **model object** is a relation between some input data objects and some output data objects. A **workflow object** is an orchestration of tools and model objects executing over a set of input data objects. **Tools** are pieces of software that perform specific functions (*e.g.* solvers, data segmentation software). Execution, storage and networking services provide assistance in transferring resources, executing models and workflows in the appropriate compute environment, as well as further downstream activities such as processing and storing of the results.

9.2 The data object

Each **VHT data object** is a digital dataset, stored and annotated according to some basic rules. The dataset must contain quantitative information on human pathophysiology, whether measured or predicted. It must be stored and curated according to the FAIR principles to be findable, accessible (possibly through authentication and authorisation), interoperable and reusable. The dataset must be annotated with sufficient metadata, which includes information on the data object type and its position in the data space (called data object pose).

9.2.1 The data object type and pose

The **Data Object Type (DOT)** is a unique identifier associated with enough information to decide if and to what extent that data object is suitable input for a DT model. This includes information on the dataset regarding its *semantics* (what the data mean), its *syntax* (in which standardised, interoperable formats the dataset is presented), and its *accessibility* (how the dataset can be retrieved). Eventually, DOTs will be selected from a list of standardised types, possibly organised in a well-structured taxonomy or ontology (*cfr.* PART 4). The list of supported DOTs might start as a folksonomy, a user-generated way of organising content, which is periodically scrutinised and consolidated into proper VHT ontologies. VHT ontologies refer to structured, formalized representations of knowledge within the VHT. These ontologies define the relationships, attributes, and classifications of the various resources and components that make up the VHT. By providing a shared vocabulary and logical framework, VHT ontologies facilitate interoperability, advanced metadata annotation, semantic searches, and consistent integration of diverse data types across the multidimensional space of the VHT.

In computer vision and robotics, the pose of an object is the combination of the object's position and orientation. Pose estimation determines a detected object's pose relative to some coordinate system.

⁹⁸ Viceconti *et al.* arXiv:2023; 2304.06678v1. <https://doi.org/10.48550/arXiv.2304.06678>

This information can then be used, for example, to allow a robot to manipulate an object or to avoid moving into the object. In the context of the VHT, the **Data Object Pose (DOP)** includes all information to define the data object in the VHT's n-dimensional reference system (data space). This information includes the information regarding the object's *position* and *scale*. The object's scale can be expressed in terms of its grain and range⁹⁹. These concepts relate to the space/time resolution of the instrumentation used to observe the data (grain), and the size/duration of the features of interest in that data (range). Ideally, this information would be captured as part of instrumentation metadata (including its calibration) and linked to the measurement, together with a protocol describing how the data was collected. The concept of grain and range as scale representation applies well to datasets that define the variation of a quantity in space and time. But since we assume, by convention, that also scalar values (such as blood biomarkers) need to be associated with a point in the reference system of the VHT, in that case, the grain represents the least significant digit of the measurement/prediction (reproducibility of the measurement, uncertainty of the prediction). In contrast, the range could represent the uncertainty of positioning in space and time for that scalar quantity.

Depending on how the VHT will develop, there might be some necessity for developing a **general data object template**, of which every possible data object type considered in the VHT (*e.g.*, imaging, wearables) is a specialisation. Some related work was done as part of the development of the so-called Multimod Application Framework, a rapid application development library for biomedical software applications¹⁰⁰.

9.2.2 Annotation and annotation services

The VHT's organisational paradigm is established by assigning a DOP to data objects in the multidimensional reference data space. Whenever a new DOT is added, it should also be provided with the transformation functions required to calculate the DOP for each data object with that DOT.

We have identified, so far, six dimensions for the data space: three spatial coordinates, time, clustering, and credibility. All this information combined will define the DOP, however, additional dimensions might be considered during the further development of the VHT.

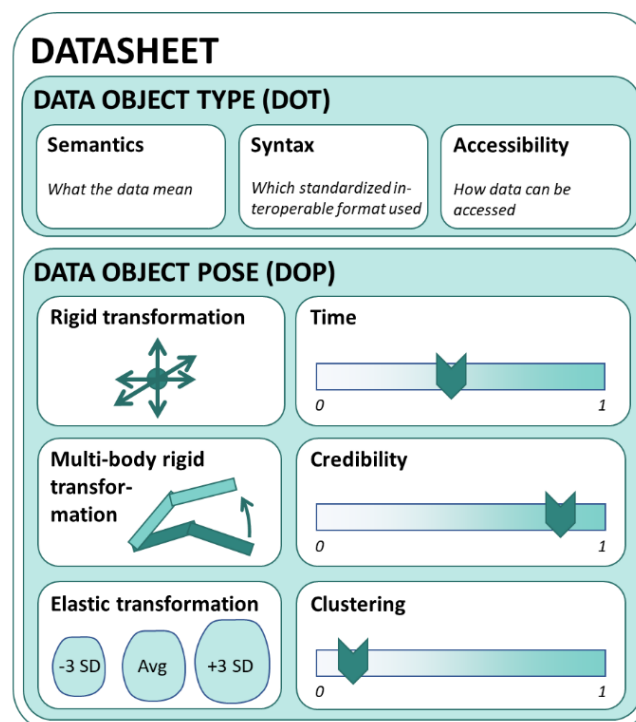


Figure 5: A graphic representation of the data sheet describing the relevant information of a data object.

⁹⁹ Bhattacharya *et al.*, PLoS One. 2021;16(5):e0251297. <https://www.doi.org/10.1371/journal.pone.0251297>.

¹⁰⁰ Viceconti M *et al.*, Proc Eighth Internatl Conf Inform Visual 2004; 15-20, <https://www.doi.org/10.1109/TV.2004.1320119>.

Dimensions 1 to 3: Space

The space is that of the human body as represented in **anatomy**. But where descriptive anatomy traditionally provides a qualitative, descriptive (semantic) representation of the spatial organisation of the body, we need a quantitative, universal representation of the anatomical space. This anatomical space is the average body geometry of all human beings, the average human body. We call this the VHT anatomical template.

Data objects can be defined over 0, 1, 2, or 3 **spatial dimensions**. For example, the systolic blood pressure of a subject is a 0D data object, how the blood flow velocity varies along the length of an artery is a 1D data object, the distribution of temperature over a region of the skin is a 2D data object, the distribution of bone mineral density in a bone is a 3D data object.

Each data object (except 0D objects) represents the spatial variation of its values **using a reference system** (implicit or explicit). So, in a 3D data object, the value corresponding to the coordinates (0, 0, 0) places such value at the origin of this reference system. In addition, each data object is referred to a specific individual and their anthropometry. But to simplify automatic annotation, clustering, and other similar operations, it is convenient that each data object is mapped to a conventional anatomical space by posing it with respect to the VHT anatomical template (which provides the DOP).

Using a conventional space requires each data object type to be provided with one or more annotation functions to transform it into the conventional spatial reference system. 3D objects can be easily posed in the **anatomical space**; the bone mineral density distribution of a patient's femur can be posed in the VHT space region corresponding to the femur of the anatomical template. With some caution, 2D and 1D objects can also be posed with respect to the anatomical template. However, 0D objects do not have an anatomical location. But because all data objects in the VHT must have one, all 0D data objects are, by convention, mapped on a 3D point located in a conventional point in the anatomical space. So, for example, the systolic blood pressure value could be posed at the centre of the heart region in the anatomical template or in the arm region where the sphygmomanometer was applied.

For spatially localised data objects, it is necessary to define **appropriate transformations** to calculate the pose of the data object in the conventional anatomical space¹⁰¹. Spatially localised data objects usually have a volume, and thus they occupy a region of the anatomical space, usually represented with the bounding box. As such they do not only have a position in space but also an orientation. Moreover, the spatial transformations are also necessary to produce the average data objects to populate the clustering axis. Because we want to preserve the integrity of the original data objects, and because applying a spatial transformation is usually not computationally expensive, it is convenient to keep the data object in its original spatial localisation and store the transformation parameters that pose it in the conventional anatomical space as metadata, that are applied on the fly to the data when required.

The **minimum set of transformation parameters** should include the roto-translation matrix that rigidly transforms from the reference system of a spatially localised data object to a pose that is anatomically appropriate in the reference system of the VHT anatomical template. Anatomically appropriate means centred and aligned. So, for example, the 3D surface of a femur of a 70 year old female patient would be transformed to have its centroid coincident and its principal inertia axes aligned with those of the femur of the VHT anatomical template. Of course, that subject's femur would generally be smaller or bigger than the average femur.

Users can also add the transformation parameters for more complex spatial transformation, such as an affine transformation that scales the data object so that its bounding box matches that of the template organ or fully elastic registrations that match the geometry of the data object to that of the corresponding template. Elastic registrations are used to generate average datasets for specific population clusters.

As mentioned before, among the essential metadata for each data object, one must include the **spatial range and grain**, which facilitates the definition of the spatial scale in which the data object is defined.

The anatomical mapping of all data objects to a conventional anatomical space poses **some challenges**. For example, how to handle datasets that refer to multiple anatomical locations (*e.g.*, recordings of a

¹⁰¹ Model, atlas or DL-based approaches are currently adopted in clinical practice to align multimodal medical imaging collected in clinical practice.

multi-lead electrocardiogram). In such cases, one could position the dataset in correspondence with the heart centre or the chest region's centre. Or, if the anatomical location of each lead is available, one could decompose the dataset into multiple data objects, one for each channel, and place them at the anatomical location of their lead. In this case, the metadata should include the information that the same experiment produced all those measurements.

Dimension 4: Time

The dimension of time may require three **different representations** depending on the use case. The first is an absolute time axis where each data object is placed at a specific date/time coordinate (for example, using the World Wide Web Consortium (W3C) date and time formats specifications). However, in some cases, it is more useful to represent time as the age, the relative time passed since birth. For this purpose, we need to add a birth date, actual or estimated, to the metadata. Last, in other use cases, data objects could be organised along a normalised time axis that spans from zero (birth) to one (death). For this, we also need in the metadata a death date, actual or estimated. Since these three representations can be easily generated on the fly when the necessary metadata is available, we recommend storing the date/time at collection, birth date and death date, leaving it to the user interface to choose the most convenient representation depending on the use case.

Time-varying datasets will range between two date/time points. However, storing the start date/time and duration might be convenient instead. Since the time of birth is not a generally available information, we will assume that all subjects born on a given day were born at 12:00 (noon) because sixty per cent of babies are born during the day, between 6 A.M. and 6 P.M.

Among the essential metadata for each data object, one must include the **temporal range and the grain** of the data object, which makes it possible to define the time scale in which the data object is defined.

Dimension 5: Clustering

Human pathophysiology varies widely between individuals. Clustering indicates the level of patient-specificity of a data set, from patient-specific data (fully un-clustered, $k = 0$) to fully clustered (homo sapiens, $k = 1$ in case only human data would be allowed in the VHT). On the one hand, the VHT should contain quantitative knowledge about the pathophysiology of many individuals; on the other hand, in several situations, such knowledge needs to be represented or is provided only as an average of a group of individuals.

To ensure irreversible anonymisation of patient-specific data, the metadata includes a unique data object ID and a unique PatientID, not associated with the individual identity. Where necessary, a LocalPatientID can be used to support pseudo-anonymisation schemes.

Each data object type must include an **averaging function that enables clustering** among its annotation services. For data objects defined in space, this is typically an elastic registration function; for time-varying objects, it might involve a synchronisation function, and for data objects not defined in space-time, these are averaging functions in the statistical sense.

When data sets are stored with sufficient metadata, this can inform the **creation of additional groups** in the data space, calculating clustering value of these groups as well as the criteria used for clustering. This means that on the Clustering axis, there might be at the same coordinate multiple data objects for the same DOT type, each obtained with different clustering criteria. For example, under $k = 0.5$, we could have a male-female, healthy-diseased, or a clustering above or below 43 years of age. How do we map clustering rules with the same k value? One possibility is using the same interface used to handle multiscale data. Suppose we navigate in the VHT infrastructure to the group of CT scans of the femoral bone: in that group, we will find clinical CTs in which the whole femur is depicted, microCTs in which only a small tissue biopsy is depicted, and nanoCTs in which only a few trabeculae are depicted. The user interface will have to represent the fact that in that group there are data objects defined on a different space-time scale and provide a way to navigate this additional organisational principle.

Dimension 6: Credibility

When a new data object is added, it is placed at the lowest **level of credibility** (non-qualified data). The data owner can submit a data object to the credibility transformation function. The higher the credibility of a data object, the higher its value. Depending on the level of credibility that the owner is requesting, the application must be informed by a smaller or greater amount of information that captures the provenance, the quality, the metrological properties (or computational credibility properties if the data are computed), and the certifications of the instrumentation/software used. For high levels of credibility, the request might be evaluated by a panel of experts, possibly in coordination with regulatory agencies. If the data are computed, the credibility of the data object will depend on the credibility of the computational model that produced it, and the uncertainty on that data should be available based on an uncertainty quantification of the computational model, taking into account the uncertainties on the input data and parameters of the computational model.

The disease state as a dimension?

Considering the main aim of the VHT is healthcare, one could wonder why there is **no disease axis** to the multidimensional space we use as an organisational paradigm. We can represent the human body as a closed system in a state defined at each instant by the values assumed by its state variables. Diagnosis, the identification of a disease, is not only characterised by symptoms (unusual combinations of values for some state variables) but also by some causal relationships. Indeed, in medicine the term syndrome indicates clusters of symptoms that might be associated with different diseases (differential diagnosis) or that lack any causal explanation. But being healthy (or diseased) is also a subjective construct of the patient. So, the concept of disease is too complex to be used as an organisational paradigm. Still, VHT users will be able to annotate their data objects with disease-describing metadata using clinical terminology dictionaries such as SNOMED-CT¹⁰², ICD¹⁰³, or full ontologies such as EBI's DOID¹⁰⁴.

9.2.3 An illustration of the data object pose and its six dimensions

We use the hypothetical generation of the VHT Anatomical Template to illustrate how these six dimensions are defined. Let us imagine having a large collection of 3D body scans of humans of all ages, genders, *etc.* Information related to the scanner, protocol, calibration procedures *etc.* are captured in the DOT. In theory, all scans were taken with the subject in the same stance (standing with the feet slightly apart, arms along the sides with the palms forward).

Each dataset is expressed with respect to an implicit reference system specific to the type of scanner used. We **position the data object** on the **time axis** in correspondence to the date/time of the scan, but we also add the subject birthdate, so we can represent the data objects also with the age axis.

Assuming the scans were all performed with fully certified 3D scanners, and were assessed to be of good quality throughout all captured regions/organs/anatomies, we place all datasets at value 1 on the **credibility axis** (which ranges from 0 for non-qualified data to 1 for fully certified measured data). Since each dataset refers to an individual, we place all of them at 0 on the **clustering axis**.

If we now select all datasets for individuals of a certain age, we can perform some spatial normalisations. The first normalisation operation assumes the body is a rigid object. We define an anatomical reference system (*e.g.*, origin in the projection of the centre of mass on the floor, X oriented from posterior to anterior, Y from medial to lateral, and Z from feet to head) and calculate for each dataset the **rigid transformation** so that they are all **aligned to the anatomical reference system**. The second normalisation operation assumes the body is a kinematic chain, *e.g.*, a set of rigid bodies articulated through idealised joints. We define in the anatomical reference system an ideal body posture. Then we calculate the **multi-body rigid transformation** for each dataset that **aligns each scan to this ideal body posture**. The third and last spatial normalisation assumes the body is an elastic object. We use statistical atlas techniques¹⁰⁵ to calculate for each time point the average body shape and then calculate the **transformation of each dataset to this average body shape**. The vector of average body shapes at different ages is the VHT anatomical template. Each new VHT data object must be posed to this anatomical template.

¹⁰² <https://www.snomed.org/>

¹⁰³ <https://www.who.int/standards/classifications/classification-of-diseases>

¹⁰⁴ <https://www.ebi.ac.uk/ols4/ontologies/doid>

¹⁰⁵ Sotiras *et al.* IEEE Transactions on Medical Imaging, 2013;32:7,1153-1190. <https://www.doi.org/10.1109/TMI.2013.2265603> .

9.3 The model object

9.3.1 The model object type and pose

For consistency of the user interface, we should consider defining also for the model objects a **Model Object Type (MOT)** and **Model Object Pose (MOP)**. The MOT must contain the list of DOTs from its input and output set. It must also contain all the metadata (following existing standards, *cfr.* PART 4) to ensure traceability, such as version, author, *etc.* Data objects are annotated with their pose in the VHT multidimensional space. Location in the anatomical space, age, and clustering does not apply to models, so model objects and workflow objects will assume a pose that is a function of their input and output data objects. Consistent with the data objects, the MOP could be automatically calculated as:

1. Average of all poses of all inputs and outputs;
2. Average of all poses of all outputs;
3. A function in the multidimensional space that links some data objects (inputs) to others (outputs).

When a new model is added to the VHT, first one needs to check if, for each input and output of the model, a valid DOT has already been defined. If this is not the case, the new DOTs need to be added before the model can be published on the VHT.

In the VHT, **model objects are defined as data space crawlers**. A VHT model requires a finite number of inputs, described in terms of DOTs and DOPs, and produces, upon successful execution, a certain number of outputs, also described in terms of DOTs and DOPs. When a model is active, every time a new data object with the necessary DOT is added to the data space, the VHT model could be automatically executed. Its outputs are also added to the data space in the appropriate DOP. This is why they are defined as data space crawlers: we can imagine model objects like little insects that crawl the honeycomb of data objects, “eat” some data objects from certain honeycomb cells and “lay” some new data objects in other cells. Thus, every time we add to the VHT a group of data objects that constitute a valid input for a model object, the dataspace will automatically be enriched with new predicted data. This implies that VHT models must execute in batch mode. However, human interaction is still possible using a “person-in-the-middle” paradigm¹⁰⁶.

9.3.2 Execution, storage and networking services

An “**eager**” **execution model** as described above, while perhaps desirable, would pose in practice a lot of challenges in terms of required resources (computation and storage). A possible solution could be this: when a new valid input is added to the VHT all models that can use that input are automatically executed. But this only means that these simulation jobs are added to a queue, where they will stay indefinitely, until someone binds one job to two additional pieces of information. The first defines the **computational resources** that can be used to run that simulation. This means a specific simulation cluster, but also a valid account in it to which sufficient computational resources are associated to run the simulation. The second defines if the results of the simulation should be **deployed** first in the sandbox of the user requesting the execution (so they can inspect the outputs before publishing them) or published directly in the VHT.

There are two other important technical aspects that need to be addressed: **remote execution and orchestration**. The VHT will run on a single computer cluster with some storage in the simplest scenario. All data objects are stored in this storage, and all model objects execute on the computer cluster. But as soon as we imagine more complex architectures, inherent to federated approaches, we might have a situation where the storage that contains the data objects, and the computer that executes the model objects, are not co-located. To ensure maximum flexibility, we can imagine a scenario where both data objects and model objects are portable, the first using data replication services and the second using container architectures. This would allow the creation of a rule-based system that decides case by case if it is better to move the data or the models. The second issue is model orchestration, which is addressed by Workflow Objects.

¹⁰⁶ Yetisgen-Yildiz *et al.*, Annual Symposium proceedings/AMIA Symposium. AMIA Symposium, 2010:1316.

9.4 Workflow objects

Complex phenomena are more easily modelled as orchestration of multiple models, each capturing a particular aspect of the knowledge available on that phenomenon. Typically, orchestrations define the data flow (how data is passed from one model to the other) and control flow (in which order and under which conditions the models execute).

An orchestration is defined as strongly coupled when the models involved need to repeatedly exchange data. These models can be orchestrated using one of several available specialised libraries¹⁰⁷. For efficiency requirements, this special class of orchestration is better exposed as a single monolithic executable, which internally invokes various codes that exchange data via memory.

In all other cases, **orchestrations can be represented with a scientific workflow**, a structured series of computational or experimental steps designed to perform a specific scientific task or analysis. Workflows do not only contain control flow and data flow instructions. They can define data selection algorithms (which of the data objects available for the input types are to be used in the simulation), pre-processing and post-processing algorithms, *etc.* In a way, a workflow can be constructed (for example using knowledge graphs) to represent the software side of a Digital Twin, where selected data is processed with selected models and services to provide the necessary output. This could be very useful in all cases where a decision does not depend only on data or models themselves, but also for what they are used for, enabling increased interoperability among models and better reproducibility across research teams.

If the control flow is a Directed Acyclic Graph (one model executing after the other), the workflow may be left implicit in the data structure, thanks to the VHT automatic execution.

AI tools are excellent candidates for performing resources orchestration and optimization. By leveraging a machine-readable overview of available resources (along with computational requirements and pricing), AI could optimize the execution of VHT computations in terms of cost and speed, enhancing efficiency and utility.

9.5 The credibility axis

9.5.1 Accuracy and credibility

When adequately annotated digital data object model objects are added to the VHT, they always start with zero credibility. As **evidence of accuracy** is added, their credibility can be increased according to rules defined by the community of practice for each type of resource.

The VHT will preferably contain quantitative information obtained from humans and provided separately for each individual. For the datasets that are semi-quantitative, are not obtained in humans, or are provided only in terms of population average, the credibility problem becomes extremely complex and cumbersome. In the first instance, the choice could be made to not provide mechanisms for credibility upgrade of the latter category of resources. In other words, non-quantitative, non-human or non-individual datasets can only stay at credibility zero.

Data objects can be of two kinds: statements (*e.g.*, surname = Smith) and measurements. For statements, accuracy is a binary concept: the surname is Smith or it is not. In some contexts, the accuracy of statements is called correctness. The accuracy of measurement is defined in a metrological sense¹⁰⁸; it is expressed by a pair of values, a measure of **trueness** and one of **precision**. Trueness expresses the systematic errors affecting the quantification; precision expresses the casual (random) errors. In metrology, the validity of a **measurement chain** (and thus of all the information produced with it) is qualified by first measuring the same specimens with a measurement chain known to be at least one order of magnitude more accurate than the one that is being tested (reference chain). Then, the series of measurements done with the chain under testing are compared with those obtained with the reference chain. This provides the accuracy of the information produced with that measurement chain. Subsequently, the same specimen is measured repeatedly: the repeated measurement's variance provides the information's precision. Of course, accuracy (trueness and precision) can vary depending

¹⁰⁷ Borgdorff *et al.*, Phil. Trans. R. Soc. 2014 ; A372, 20130407. <https://www.doi.org/10.1098/rsta.2013.0407>

¹⁰⁸ <https://doi.org/10.1007/s00769-006-0191-z>

on the measured value. Thus, a metrological campaign is conducted over various specimens where the measured value ranges between a minimum and maximum, which are considered the interval of validity of that measurement chain; accuracy and precision are provided as the average values over these repeated assessments. Thus, the credibility of quantitative information is fully characterised by identifying which measurement chain was used to generate it and the average precision and accuracy of this measurement chain over its validity interval.

The **credibility of predictive models**, is expressed in terms of predictive accuracy against controlled experiments. But it also requires the decomposition of the predictive error in *numerical, aleatoric and epistemic* components¹⁰⁹ through the process known as verification, validation and uncertainty quantification (VVUQ, see part 4). However, in the most common case where numerical errors are negligible compared to the others, validation can estimate trueness and uncertainty quantification precision.

9.5.2 Quantifying credibility

Thus, a more helpful categorisation for the credibility axis could be if we divide **data** into *categorical* and *quantitative*. For categorical data, credibility can be either 0 or 1, depending on if they have been checked or not. Credibility is expressed in terms of trueness and precision, whether the value is measured or predicted. But because the concept of credibility is not independent of the context of use (CoU), it is recommended that even for quantitative data, the credibility axis admits only a few conventional values. Below is a first proposal of such credibility values.

- 0.00 – no credibility;
- 0.25 – partial evidence of credibility with respect to the CoU is provided;
- 0.50 – complete evidence of credibility with respect to the CoU is provided;
- 0.75 – evidence of credibility is considered sufficient with respect to the CoU by the Community of Practice;
- 1.00 – credibility is certified with respect to the CoU by a third party (such as a regulatory agency).

When categorical data is obtained from the interpretation of measurements (*e.g.*, medical images, ECGs), the credibility assessment is a hybrid between both systems. Has the interpretation been done by an expert or a panel of experts? And did all experts agree? The answers to these questions will lead to a credibility between 0 and 1.

For the **model** resources, assuming that the *predictive error* is due to the sum of the numerical, aleatoric, and epistemic errors, the concepts of trueness and precision can be defined for models as established for the data resources. Not all models are equation-based or use numerical methods to solve equations, and for those models the numerical error is zero. Otherwise, it must be quantified. The goal is to demonstrate that the numerical error is negligible compared to the sum of the other two such that its effect can be neglected. Thus, the most helpful information is the upper boundary of the numerical error. Once this verification is done, we can use experimental validation to quantify the accuracy and uncertainty quantification to quantify the precision. Same as for the data resources, model resources should be annotated with some limits of validity: these can be of epistemic origin (the knowledge used to build the model carries some limits of validity). Still, we think it is more helpful to indicate the limit of validity as the range of input values explored in the validation experiments to provide a basis for assessing the model's applicability in the ASME VV-40:2018 sense.

Credibility of **data-driven (AI) models** is the subject of ongoing discussions and consultations by regulatory agencies¹¹⁰. Elements important to assess credibility of such models within their context of use include the model architecture and parameters, the data used to develop the model (relevance, reliability), model training, potential biases etc. A European **AI Testing and Experimentation Facilities** (TEFs), TEF-Health¹¹¹, assists AI developers (and assessors) in the health space by providing state-of-the-art testing environments for validating trustworthy AI technologies and accelerating their adoption in the market.

¹⁰⁹ Viceconti *et al.*, IEEE J Biomed Health Inform. 2020;24(1):4-13. <https://www.doi.org/10.1109/JBHI.2019.2949888>.

¹¹⁰ <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological>

¹¹¹ <https://tefhealth.eu/home>

Traditionally, **uncertainty quantification** aims to explore how the aleatoric uncertainty affecting the inputs propagates into the outputs. However, if we plan to use uncertainty quantification as a measure of precision, we should also include whatever residual epistemic uncertainty affects our model in the uncertainty quantification. The most concrete example is to include the uncertainty due to using some population averages as inputs instead of subject-specific values. But it also applies to environmental factors, *etc.*

10 Data and data generating hardware for VHT

Data generation in advanced healthcare technologies originates from diverse sources, including omics¹¹², multi-modal sensors in wearables and implantables^{113,114}, nanomedicine¹¹⁵, (advanced) imaging¹¹⁶, body-on-chip models¹¹⁷, and other enabling technologies. These diverse approaches can collect data to inform and feed disease and organ models. In the context of the VHT, developing and providing advanced tools, technologies and infrastructure to accelerate the generation of (new) deep data remains an important endeavour, enhancing our knowledge of the means to support prevention and personalised treatment of diseases. These can include, but are not limited to, engineering advances in energy-efficient sensing, computing, communication and cloud/fog/edge technology together with adequate modelling and computational approaches, including AI, to generate and exploit big and deep data.

10.1 Data Generation Sources and Sensing Technologies

Advanced imaging technologies have revolutionized the ability to visualize organs and biological processes with unprecedented detail and precision. Innovations such as high-resolution 3D imaging, functional MRI (fMRI), and molecular imaging enable clinicians and researchers to observe structural and functional aspects of tissues in real time and VHT with incredible high accuracy. Imaging techniques reveal cellular and subcellular dynamics, offering deep insights into disease progression and therapeutic responses are of future interest. Artificial intelligence (AI) further amplifies the potential of advanced imaging by enhancing image analysis, segmentation, and pattern recognition. AI algorithms can identify subtle anomalies, predict disease trajectories, and integrate imaging data with other patient metrics for more comprehensive diagnostics.

Multi-modal biosensors¹¹⁸ represent a transformative leap in health monitoring, integrating the ability to track complex, dynamic and highly diluted health analytes in body fluids alongside traditional vital signs like heart rate, respiration, and blood pressure. These systems, encompassing wearable and implantable embodiments, are designed to simultaneously measure multiple biomarkers, such as glucose levels, stress hormones, oxygen saturation, and metabolites, offering a comprehensive view of a patient's physiological state. Advanced technologies enable these devices to detect subtle biochemical changes in real time, significantly enhancing early disease detection and personalized intervention. The integration of multi-modal biosensors with digital health platforms further amplifies their potential, allowing seamless data integration, real-time analytics, and remote patient monitoring.

Advances in miniaturization and sensor processing have led to battery-operated, highly efficient devices that provide continuous health monitoring through point-of-care and wearable sensing systems. Their small form factors make them adaptable for everyday health tracking^{119,120} and chronic disease management. These systems integrate signal processing within the sensor, enhancing portability and reducing reliance on external laboratory devices. **Wearables** are real-time data generators, capable of supporting remote patient monitoring and mass screening of populations for disease detection, prediction and progression monitoring. In this context, the value of wearable technologies is based on real-time data collection, non-invasiveness and the ability to support DT representation with miniaturized devices. The majority of technological efforts to date have focused on physiological monitoring of vital signs¹²¹ and of activity¹²² (*e.g.*, cardiac monitoring, core body temperature, blood pressure, step count *etc.*), using personal devices (*e.g.*, smart watches, smart rings). However, there is an urgent need for advanced monitoring technologies for specific biomarkers that take into account the

¹¹² <https://hsrc.himmelfarb.gwu.edu/gwhpubs/5315>

¹¹³ <https://doi.org/10.1371/journal.pdig.0000104>

¹¹⁴ <https://doi.org/10.3390/s110605561>

¹¹⁵ <https://doi.org/10.1038/s41568-022-00496-9>

¹¹⁶ <https://doi.org/10.1038/s41583-021-00441-z>

¹¹⁷ <https://doi.org/10.1038/s41578-018-0034-7>

¹¹⁸ <https://doi.org/10.1002/adfm.202403122>

¹¹⁹ <https://doi.org/10.1038/s41578-022-00460-x>

¹²⁰ <https://doi.org/10.1038/s41528-021-00107-x>

¹²¹ <http://dx.doi.org/10.1136/bmjinnov-2019-000354>

¹²² <https://doi.org/10.3390/s21051802>

most relevant dynamics of disease symptoms and physiological parameters, such as, for instance, the detection of cytokine storm in COVID-19 or sepsis patients¹²³. The VHT platform should be designed to exploit such real-time data as input to future generations of DTs.

Interest in **point-of-care** and large-scale bio-chemical sensing has triggered a lot of research in the past decade into highly scalable technologies that could benefit from the scaling and low power consumption of semiconductors chips (*e.g.* ISFETs¹²⁴ and CHEMFETs¹²⁵). Nevertheless, numerous challenges remain related to such devices aiming at ion, metabolite, protein and hormone sensing in biofluids as well as to their readout electronics, in terms of sensitivity into sub-nM range, selectivity and integrated multi-parameter sensing. Recent years have seen an increased research effort in semiconductor sensors benefitting from new 1D and 2D semiconductor nanostructures¹²⁶, as well as new readout architectures and topologies.

Finally, another important aspect of advancing the state-of-the-art in wearable devices is the **biofluid** that is used for non-invasive or minimally invasive approaches - as compared to blood, the medical gold standard. While sweat¹²⁷ appears in many recent reports as fluid of interest for non-invasive sensing, the correlations of markers in sweat with blood are poorly understood and contamination remains an issue. Despite a semi-invasive access, interstitial fluid appears to hold more promise for biosensing¹²⁸. Additionally, there is ongoing research in breath and saliva analysis with the target of replacing invasive blood tests with breath/saliva biosensors, but these are still at an early stage.

Implantable medical devices have evolved for data generation beyond traditional applications like pacemakers and cardiac stimulators to encompass sophisticated, smart biodegradable systems¹²⁹, capable of transformative clinical impact. Modern implantable devices integrate closed-loop systems for real-time monitoring and intervention, such as smart drug delivery platforms for advanced cancer treatments that release therapeutics in response to precise biomarkers, enhancing efficacy while minimizing side effects. Similarly, implantable electrodes for neural stimulation offer groundbreaking solutions for conditions such as Parkinson's disease, epilepsy, and spinal cord injuries, supporting data generation for adjustable therapies. An exciting recent innovation is formed by transient implantable sensory systems, with a predefined operational lifespan that dissolve safely within the body after completing their function. These bioresorbable sensors eliminate the surgical removal, reducing risks of infection, scarring, and patient discomfort. Clinically, they are valuable in applications such as building future VHTs for post-surgical predictive monitoring, where they provide critical data on healing and complications, or in controlled drug delivery, where they ensure precise dosage over time.

Nanotechnologies enable precise data generation at the molecular level, offering real-time insights into disease mechanisms and treatment responses. Nanomedicine plays a key role in personalized healthcare strategies, particularly for targeted drug delivery and localized monitoring of conditions like cancer.

Advanced high-throughput omics technologies enable simultaneous examination of genomes, transcriptomes, proteomes, and metabolomes, generating vast datasets critical for Digital Twins. Innovations in nanoscale CMOS technologies have advanced genomics¹³⁰ by enabling compact, high-performance biosensors and sequencing devices. These CMOS-based chips have reduced the cost and time for genome sequencing, making it more accessible for research and clinical applications.

Microphysiological systems such as organoids and body-on-chip technologies¹³¹ utilizing microfluidic systems embedded with sensors to simulate the functions of human organs on a miniature scale can generate extremely insightful data for building VHTs. These systems recreate complex physiological

¹²³ <https://doi.org/10.1038/s41392-021-00679-0>

¹²⁴ <https://doi.org/10.3390/s90907111>

¹²⁵ <https://doi.org/10.1109/TED.2022.3144108>

¹²⁶ <https://doi.org/10.1021/acsnano.9b03632>

¹²⁷ <https://doi.org/10.1159/000504387>

¹²⁸ <https://doi.org/10.1038/s41551-020-00679-5>

¹²⁹ <https://doi.org/10.1038/s41578-019-0150-z>

¹³⁰ Heather, JM, Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics*. 2016 Jan;107(1):1-8. doi: 10.1016/j.ygeno.2015.11.003 . Epub 2015 Nov 10. PMID: 26554401; PMCID: PMC4727787.

¹³¹ <https://doi.org/10.1038/s41576-022-00466-9>

environments by channelling fluids and mimicking tissue interactions, allowing researchers to study disease mechanisms, drug responses, and organ-specific functions in a controlled, high-throughput setting. The integration of real-time sensors enables continuous monitoring of biomarkers, providing invaluable data for personalized medicine and improving the accuracy of preclinical drug testing. This innovative approach bridges the gap between traditional cell cultures and animal models, offering a more precise, ethical, and scalable platform for medical research and development.

Exposome science and sensing integrates environmental data with biological responses to capture the full impact of external factors like air quality, food, water, and lifestyle on health. The molecular phenotypes underlying environmental effects on health are important yet still poorly understood¹³². Sensors monitor pollutants, contaminants, and personal behaviours, offering valuable insights for VHTs that assess how these factors interact with individual biology.

Box 12: Success story – scDrugPrio

Success story: scDrugPrio - using single-cell transcriptomics for personalised drug selection

Status: ready for clinical testing

Website: <https://github.com/SDTC-CPMed/scDrugPrio>

The Swedish Digital Twin Consortium published a vision paper on the construction of DTs of individual patients based on combining clinical routine data with multiomics data down to the single cell level. Such DTs are constructed based on computational network models of thousands of disease-relevant variables. These twins can then be computationally treated with thousands of drugs to find the optimal drug or drugs for the patient¹³³.

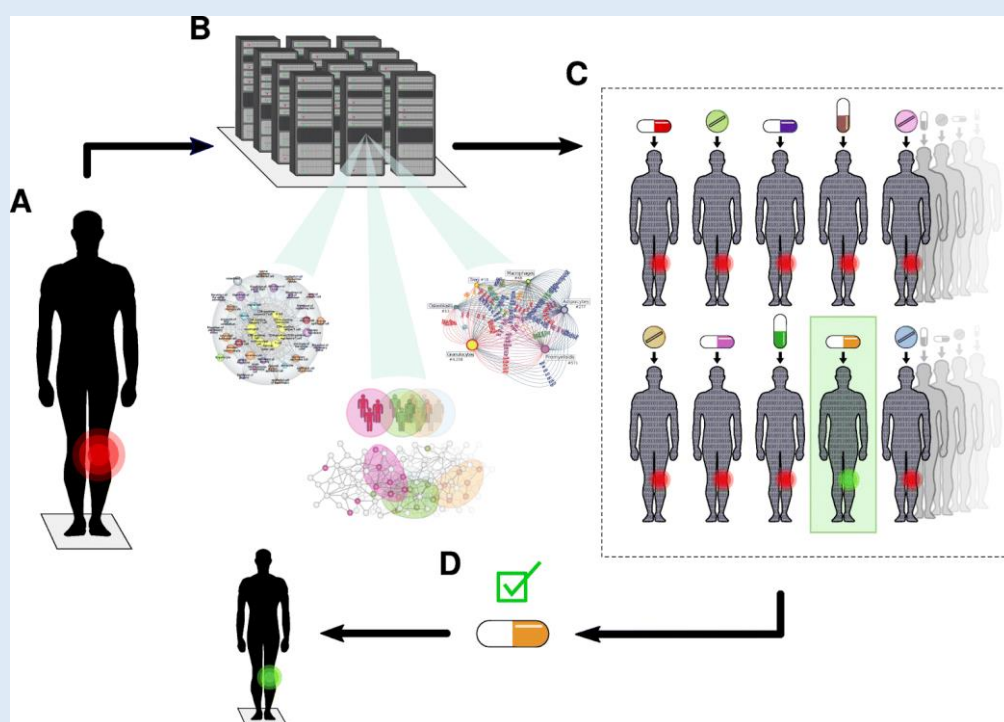


Figure 6: The Digital Twin concept for personalised medicine. A. An individual patient has a local sign of disease (red). B. A Digital Twin of this patient is constructed in unlimited copies, based on computational network models of thousands of disease-relevant variables. C. Each twin is computationally treated with one or more of the thousands of drugs. This results in a digital cure of one patient (green). D. The drug that has the best effect on the Digital Twin is selected for treatment of the patient.

A method to apply this concept for personalised medicine, scDrugPrio, has been developed and validated in mouse model of arthritis¹³⁴. A multicellular network model of the inflamed tissue was constructed based on single cell RNA sequencing data. Next, network tools were applied to match the model with thousands of drugs.

¹³² <https://doi.org/10.1038/s41467-022-34422-2>

¹³³ <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-019-0701-3>

¹³⁴ <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-024-01314-7>

The top-ranking drugs were validated by extensive *in vivo* (mouse) and *in vitro* studies. After validation, scDrugPrio was translated to human patients with inflammatory diseases. An important difference between the mouse model and human diseases was the great cellular and molecular heterogeneity between patients with the same disease, which resulted in different drug predictions. These findings simultaneously show the potential of DTs in personalised drug predictions in general, and the need for high-resolution DTs in order to make accurate personalised drug predictions.

10.2 Advanced Hardware Architectures and Hardware-Software Codesign

To process the vast data generated for and by DTs, advanced hardware architectures and hardware-software co-design are essential. These innovations optimize performance, reduce power consumption, and ensure scalability for personalized healthcare applications.

Low-power sensors^{135,136} are crucial for continuous monitoring in Digital Twin systems. Advances in **energy-efficient** microelectronics and energy harvesting techniques allow for long-term, real-time data collection without frequent recharging, ensuring seamless operation in wearable and implantable devices. **In-sensor computing**¹³⁷ reduces the need for data transmission by enabling localized, real-time analysis. **Neuromorphic computing**, inspired by the human brain's architecture, allows for low-power, efficient processing, making it ideal for tasks like monitoring neural activity or organ function.

Advances in **wireless technologies**, such as 5G, facilitate faster, more reliable data transmission between sensors and central systems. This enables real-time data exchange for VHTs, crucial in applications like chronic disease management and emergency care, where timely analysis is essential.

Cloud computing provides scalable storage and analysis for patient data, while **fog computing** processes data locally, reducing latency and bandwidth usage. Together, these technologies enhance the VHT ecosystem by enabling near-instantaneous decision-making in healthcare settings. **Edge AI** integrates artificial intelligence capabilities directly into the sensor or edge device¹³⁸. This local data processing enables real-time analysis and immediate responses¹³⁹. This approach offers real-time event detection in clinical settings and supports autonomous systems focused on inference rather than training. This allows for rapid interventions in medical scenarios, reducing latency, enhancing privacy, and making personalized healthcare more proactive and adaptive. As healthcare shifts toward a more proactive model, Digital Twins^{140,141} represent a significant leap in our ability to tailor interventions to individual needs, ultimately enhancing patient outcomes and overall healthcare efficiency. In the future, **Edge-to-Cloud integration** will optimize specific tasks of sensing, processing and communications, when deploying Digital Twins for healthcare.

Box 13: Success story – DIGIPREDICT project

Success story: DIGIPREDICT – revolutionizing patient monitoring with smart biopatches and painless fluid collection

Status: clinical testing

Website: <https://www.digipredict.eu/>

The DIGIPREDICT project has advanced the frontier of patient monitoring by developing an innovative platform that integrates multimodal biosensors, painless microneedle technology, and AI-driven analytics into a smart biopatch. This wearable solution enables real-time monitoring of critical biomarkers in hospital settings, offering a transformative approach to early disease detection and personalized care.

Building on the success of interstitial fluid (ISF) glucose monitoring technologies like Abbott's FreeStyle Libre, DIGIPREDICT extends ISF sensing to a broader range of analytes. The smart biopatch incorporates state-of-the-art sensors developed by EPFL based on its collaboration with Xsensio SA that measure lactate,

¹³⁵ <https://doi.org/10.3390/s21051802>

¹³⁶ <http://doi.org/10.1021/acsnano.8b07413>

¹³⁷ Mao, L., Neuromorphic Sensing: A New Breed of Intelligent Sensors ACS Sens. 2023, 8, 8, 2896–2897.

¹³⁸ AI at the edge - a roadmap - White Paper Artificial Intelligence: Jan M. Rabaey, Marian Verhelst, Jo De Boeck, C. Enz, Kristiaan De Greve, Adrian M. Ionescu, Myunhee Na, Kathleen Philips (editors), IMEC, 2020.

¹³⁹ <http://doi.org/10.1111/cts.12884> ..

¹⁴⁰ <https://doi.org/10.3390/jpm1411101>

¹⁴¹ <https://doi.org/10.1038/s41746-024-01073-0>

pH, and C-Reactive Protein (CRP), all in near-real time. This marks a significant evolution in wearable diagnostics, providing deeper insights into metabolic and inflammatory states critical for early intervention in conditions such as sepsis and cardiovascular dysfunction. It includes: (1) lactate monitoring: an ultra-scaled lactate sensor uses amperometric measurements to continuously track lactate concentrations in ISF, (2) pH sensing: a pH sensor based on silicon junction-less nanowire transistors offers precise measurements, essential for monitoring acid-base balance in patients with critical illnesses, and, (3) CRP Detection: a novel CRP Silicon-On-Insulator FET sensor detects inflammatory responses at point-of-care, enabling early identification of systemic inflammation.

Central to this innovation is the integration of MEMS-based microneedles developed by Ascilion. These ultra-sharp, hollow microneedles provide painless, continuous access to ISF, eliminating the need for invasive blood draws. Each chip, containing an array of microneedles, efficiently channels fluid to the biosensors via a capillary system, ensuring accurate and consistent sampling. This approach enhances patient comfort and enables continuous monitoring without disrupting clinical workflows.

Data collected from the biopatch is processed at the edge using advanced AI algorithms embedded within the device. This enables real-time analysis of biomarker dynamics, such as the early detection of cytokine storms. The system integrates with a smartphone application, providing clinicians with predictive insights and alerts. Machine learning models enhance accuracy and support personalized, data-driven interventions. In addition, the DIGIPREDICT Digital Twins use organ-on-chip technology, integrating human stem cell-derived tissues within dynamic environments that simulate physiological conditions, to select the right biomarker combination for generating an accurate picture of how the disease is progressing in a patient and how well the chosen treatments are working.

The DIGIPREDICT platform has undergone rigorous validation in clinical trials at Charité in Germany and Insel Hospital in Switzerland. These trials demonstrated the biopatch's ability to monitor key physiological parameters, including blood oxygen levels, breathing rate, and critical biomarkers associated with cardiovascular and immune responses. The patch's low-power design and wireless data transmission ensure seamless integration into hospital settings, improving workflow efficiency and patient outcomes.

10.3 Integrating Diverse Data Sources

Data can be broadly categorized into three groups: clinical-grade data, research-grade data, and model-generated data, each playing a unique role in creating personalized, real-time healthcare insights. Ultimately, the success of the VHT relies on the seamless integration of these three data categories – each of which brings its unique strengths and contributions to the DT.

10.3.1 Clinical grade data

Clinical grade data provides the foundation for VHT applications, ensuring accurate predictions and simulations based on health measurements. It enables precise assessments of disease progression and treatment outcomes. The category encompasses data that is already being used as part of the healthcare and clinical practice and/or present in health-homologated repositories. This data type is expected to be generated by clinical grade technologies and their validity for the intended clinical use is ensured by the healthcare system. They do not require validation or justification to be included in the development and validation of the DT as long as the data are used with the same CoU for clinical practice (use for other CoUs requires additional validation). Hence, the quality of the data, data format, and interoperability should be addressed in the VHT but the data type validity is not questioned. Examples are clinical biochemistry tests, multi-modality medical imaging, Electronic Health Records, vital signs (blood pressure, ECG, core body temperature), clinical genetic tests, biomarkers specific to medical conditions and to disease progression/regression, *etc.* This also includes data from clinical research studies (clinical trials) and data in registries.

10.3.2 Research grade data

Research grade data enriches the VHT by providing a broader understanding of disease mechanisms, biomarkers, and potential therapeutic targets, supporting personalized medicine and predictive modelling. This category encompasses sufficiently advanced research grade technologies that are accepted in research practice across the globe. These technologies are expected to satisfy research community standards. Examples of such technologies are transcriptomics, proteomics, metabolomics, wearables with a combination of vital signs (heart rate, body temperature, SpO₂, etc), activity tracking, emerging dynamic fingerprints of biomarkers in human biofluids detected by multimodal sensors in real time), implantable technologies and organ-on-chip technology (at long term, emerging in a body-

on-chip technology). Additionally, there is personal data generated by an individual, including social media (*e.g.* Instagram, Facebook), location information, diaries, wellness apps (*e.g.* Headspace, Strava). Another important category of data consists in exposome data (such as food, air, water but also other factors resulting from lifestyle and associated stress). Data types in this data category require evaluation for their fit for purpose and raise important challenges in terms of very heterogeneous time-scale, format and interoperability.

10.3.3 Data generated or transformed by models

Data generated or transformed by models, translating clinical and research data into actionable insights. This category includes the results of models and simulations that are either transforming or generating data and creating a new data type. This category of data needs to be evaluated as their utility depends on their scope, sensitivity and accuracy which are affected by the model's input data types and model's precision, accuracy and credibility. Examples of these data types would be internal cellular or physiological parameters that could not be measured *in vivo*, *in silico* trial data, simulation results of a personalised DT, synthetic data generated with AI methodologies¹⁴², *etc.*

10.4 Data use and reuse

The practice of data reuse is a fundamental element as DT models are envisioned to be integrated at input-output level towards creation of the VHT. Hence, the methodology to be followed involves not only **integrating varied data categories** - encompassing clinical biochemistry, radiological modalities such as MRI and CT scans, genomics, proteomics, and metabolomics - but also **amalgamating cohort-specific data and data specific to individual patients or citizens** (contingent on their informed consent) as well as **simulation outputs created within the VHT platform**. The establishment of such a robust, multidimensional data catalogue is of key importance, capable of propelling the innovative potential of DT towards the creation of the VHT. A comprehensive integration of collected data can already constitute a subject-specific digital model if put into a common reference. Collection of subject-specific models then constitutes a collection/cohort that can further be used for population stratification, patient diagnosis and prognosis.

Data reuse involves strict legal and ethical considerations, especially for clinical data. Authorized and secured (anonymized) access to clinical data is an important issue and usually takes considerable time and effort to establish. Therefore, the VHT infrastructure should facilitate connections to health-homologated repositories (for a given CoU). **Strategic collaboration with other European data initiatives** such as OpenAIRE serves as an instrumental aspect of this approach. OpenAIRE, a strong proponent of the open science paradigm, contributes a multitude of invaluable data sets. The VHT platform must be designed to harmoniously align with the ethos and ambitions of GAIA-X, which advocates for a protected, federated infrastructure that enhances data availability. Similarly, the European Open Science Cloud (EOSC) reflects the VHT's ethos for endorsing data sharing and reuse within the scientific fraternity. Through these collaborations, the VHT infrastructure will augment its data assets further and stimulate a culture of open, cooperative investigation and will be an integral component of Europe's formidable digital trajectory.

Acknowledging the criticality of a harmonised and interoperable health data environment, the VHT resonates with the objectives of the **European Health Data Space (EHDS)**, which aims to elevate the accessibility, quality, and application of health data across EU jurisdictions. However, in recognizing the considerable power this confers, we also acknowledge the accompanying accountability. Our unequivocal commitment to data reuse within the VHT operations strictly adheres to European ethical and societal principles. Rigorous compliance with legal regulations and guidelines is an absolute mandate. VHT processes need to safeguard individual data privacy rights, consent and staunchly uphold principles of data minimization and purpose limitation.

¹⁴² <https://synthema.eu/>

10.5 Data transformation services: harmonising and transcending boundaries

The VHT is to be a robust, consolidated and interoperable environment that embraces diversity at both individual and population levels. It will encompass a vast variety of *data types* covering spatial and temporal dimensions of personal and population level data as well as various *data formats* within each data type. Hence, we envision the design and development of (AI-based) data transformation services within the VHT platform to seamlessly transform, harmonize, reformat and integrate various data sources (with each other or with models), helping the users to generate rich, actionable insights. **Data transformation services** produce novel data that is stored within the platform with full provenance information. As such, data transformation services promote a harmonised, interoperable platform that utilises health data to its fullest potential, irrespective of its origin, format, or intended scale of use.

10.5.1 Unit Conversion

Within the rich and varied health data landscape, diverse data sources often present information in different units or measurement systems. Data transformation services are designed to employ standardised techniques to reconcile these disparities, mapping data from one unit or measurement system to another. This harmonisation process allows data captured under specific protocols to be transformed into universally comprehensible units, facilitating seamless integration and interpretation across various models and applications.

10.5.2 Format Conversion

Complementing the unit conversion/transformation process is the ability to handle data format conversion. Users can convert data captured using one technology into a format compatible with other technologies, without compromising the spatio-temporal scales. This conversion is essential to ensuring compatibility and interoperability across the platform, allowing the most effective use of diverse data sources.

10.5.3 Dividing individuals in groups

Another important transformation service would be in segmentation of personal health data. This involves dividing the health-related information of individuals into distinct groups based on specific criteria. In this type of transformation we do not change the data object but its grouping. Some examples are listed below.

- **Medical conditions:** dividing individuals based on the specific medical conditions they have, such as diabetes, hypertension, asthma, or cancer. This segmentation is crucial in parameterizing models with the correct individual's data for understanding disease prevalence, treatment effectiveness, and tailored healthcare interventions in a Digital Twins context.
- **Age groups:** dividing individuals into different age groups, such as paediatric, adult, or elderly populations allow the models to treat these segments of the population separately. This grouping is important as health needs, risks, and treatment options may vary across different age groups.
- **Sex & gender:** grouping individuals based on their sex/gender, as certain health conditions or treatments may be specific to a particular sex/gender. For example, breast cancer models are more relevant in females and prostate issues in males. This grouping can also be useful to try to balance or compare the ratio of male vs female in *in silico* trials, to identify trends that might help address imbalance in diagnostic or treatment, or to address potential bias in *e.g.* machine learning training datasets.
- **Genetic factors:** segmentation based on genetic information, such as specific gene mutations or variations. This segmentation can help model the impact of the individuals at higher risk for certain diseases or guide personalised treatments.
- **Lifestyle factors:** segmenting individuals based on lifestyle factors like smoking status, physical activity levels, diet patterns, or alcohol consumption. This segmentation can assist in assessing (including or excluding) health risks, modelling their effects and designing targeted interventions.
- **Electronic Health Record (EHR) data:** segmenting individuals based on their medical history, including hospital visits, medication usage, lab test results, or surgical procedures. This segmentation allows models to know which data is available from the healthcare providers and can create insights into patient populations, disease progression, and treatment outcomes.
- **Treatment response:** segmenting individuals based on their response to specific treatments or interventions. This segmentation will support models for accurately predicting possible trajectories of disease progression, prognosing patient response to therapy and evaluating treatment effectiveness, identifying subgroups that benefit the most, and refining treatment plans on a personalised setting.

10.5.4 Personal to Population Level Transitions

In the context of the VHT, an important data transformation service provides the ability to fluidly navigate between individual (clustering = 0) and population-level (clustering = 1) data. By transforming and aggregating data from individual profiles, valuable insights can be generated at the population level which are stored as population level aggregates and distributions. Conversely, users can draw on these broader patterns to inform and refine their individual-level predictions and interventions when individual level data is scarce or lacking. This dynamic interaction between micro and macro perspectives drives a comprehensive approach to healthcare that respects individual uniqueness while maintaining a population-oriented outlook.

10.5.5 Interlinking Models: The Input-Output Perspective

Beyond data conversion, data transformation services should strive to facilitate the smooth interaction between different models. This involves effectively handling the input and output data of diverse models to enhance their synergistic functioning. AI can be powerful tools to facilitate the continuous updating of DTs and optimizing their orchestration. These tools could support researchers by automating the identification of appropriate DTs, optimizing computational workflows, and enabling seamless integration of diverse data and models. By harmonizing data formats, supporting standardization efforts, and automating resource orchestration, AI has the potential to streamline development of holistic, reliable predictions and simulations, and improve its cost-effectiveness and scalability.

11 *In silico* models for VHT

11.1 State of the art

11.1.1 *In silico* medicine

In general, a predictive model of human (patho)physiology is built on observations (data) with varying amounts of prior mechanistic facts or hypotheses, hereafter referred to as "knowledge". Data-driven models are those that contain no mechanistic knowledge, and mechanistic models those that are ideally based entirely on mechanistic knowledge. However, these are the extremes of the *in silico* spectrum, with most models somewhere in between¹⁴³. Lately, the boundaries have been further blurring with explainable Artificial Intelligence (AI), physics informed Machine Learning (ML) or ML-based surrogates of mechanistic models, all emerging as techniques that combine aspects of both worlds. Which modelling approach is preferred in a given situation depends on the question of interest, the Context of Use (CoU) of the model, and the available data and knowledge. The model resources in the VHT will be covering the entire **spectrum from data-driven to knowledge-driven**.

In silico medicine has achieved substantial progress in the last decade. As can be appreciated from the use cases that are illustrating the different parts of this roadmap, digital twins of single organs are being introduced in **the market**, and **clinical research and regulatory science** are in full development with several ASME and ISO **standards** in place.

11.1.2 *Data-driven models in health and care*

AI-based Digital Twins are penetrating the market at record speed, due to their ability to identify relationships in large amounts of data that were previously unrecognized as well as their ability to speed up data processing, and are the subject of many policy and strategy initiatives (*cfr.* PART 4). **AI technologies can be employed in a variety of ways to build the VHT**. AI can help **increase** the sheer number of systems that can be assessed, from *e.g.*, organ models based on image processing to the assessment of numbers of candidate compounds that may be used for drug treatments, and for customised medical devices. Related to this, AI models can help **speed up** some operator-intensive tasks within the procedure, for example, replacing manual image segmentation with automatic segmentation (*e.g.*, by convolutional neural networks), reducing variability between operators and greatly reducing model development timescales (crucial for realistic clinical uptake and use). AI technologies can be used in the **development** and incorporation of surrogate models or ML-based Partial Differential Equation (PDE) solvers, which can replace some features of multiscale models, eliminating some of the computation-intensive aspects of these simulations. Such models will also play a key role in **uncertainty quantification** studies, which are essential for implementing and adopting actionable models that can be used in *e.g.*, clinical decision support. AI can help to **parametrise and personalise knowledge-driven models**, *e.g.*, using advanced Bayesian inverse uncertainty quantification techniques and relying on increasing and enriching multimodal individual health and clinical data.

Application of AI can lead to the **generation of completely data-driven predictive models** that would *e.g.*, provide decision support on treatment options, or patient specific prognosis prediction. To allow transition into clinical practice, such models need to be transparent and explainable, *e.g.*, by integrating prior mechanistic knowledge. **Explainable AI (XAI)**, also known as Interpretable AI, is one of the key requirements for implementing responsible AI approaches. XAI enables humans to understand the reasoning behind decisions or predictions made by the AI, in contrast with the "black box" concept in machine learning, where even the AI's designers encounter challenges to explain the decision making process of the system. Finally, AI could assist in the generation of hypotheses from data, that may inform and guide the development of mechanistic multilevel models or experiments.

Next-generation Edge AI systems for healthcare are evolving towards hybrid systems such as coarse-grained reconfigurable arrays (CGRAs), for a flexible and energy-efficient acceleration of a wide range of AI-based embedded bio-signals processing kernels. **Federated learning** enables the training of ML

¹⁴³ WHITE PAPER: the role of Artificial Intelligence within *in silico* medicine (2022). <https://doi.org/10.5281/zenodo.8064147>

models for DT over distributed training data while respecting privacy, without the training data ever leaving its device of origin¹⁴⁴.

Additionally, AI, particularly Natural Language Processing (NLP) utilising Large Language Models (LLMs) will be instrumental in the identification, collection and organisation of resources for the VHT ecosystem. NLP can perform **automatic knowledge retrieval** from the vast and exponentially growing DT literature. This can take the form of NLP models answering specific user queries or generating desired thematic summaries. Additionally, NLP can be used to organise DT literature through supplementing existing **knowledge graphs** or building new ones. They may also be used within the VHT itself to guide the VHT platform users on resources that can be coupled, ways to couple them, as well as to identify specific data needs that would lead to a more comprehensive DT. Finally, NLP can streamline the communication of platform outcomes by automatically generating reports containing the necessary technical or clinical information for a specific user, such as a clinician. These aspects are discussed in more detail in the next chapter.

A major challenge of AI-based methods especially in healthcare are biases in the data, such as underrepresentation or missing data, and data quality such as highly heterogeneous data sets resulting in biased and poorly performing AI systems.

Box 14: Success story – Radiomics Digital Twins using machine learning for cancer diagnosis.

Success story: Radiomics Digital Twins using machine learning for cancer diagnosis

Status: TRL 5

Website: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11535140/>

This story describes enhancing the accuracy of diagnosing clinically significant prostate cancer (csPCa) using radiomic Digital Twins based on bi-parametric MRI (bpMRI), incorporating T2-weighted (T2w) images and apparent diffusion coefficient (ADC) maps. These twins aim to identify imaging biomarkers that reflect tumor heterogeneity and biological characteristics. A variety of feature selection strategies, machine learning (ML) classifiers, and imaging sources were employed to develop robust and generalizable predictive models.

Two datasets were used: ProstateNET (465 patients) and ProstateX2 (204 patients), with csPCa defined as a Gleason grade ≥ 2 . Preprocessing steps included bias field correction, resampling, intensity normalization, and histogram discretization. Radiomic features (1246 per sequence) covering shape, first-order, and texture-based characteristics were extracted. Feature selection methods included (i) **Filter Methods:** e.g., mRMRe and ReliefF, (ii) **Wrapper Methods:** e.g., Boruta and Recursive Feature Elimination (RFE), and (iii) **Embedded Methods:** e.g., L1-lasso and Random Forest variable importance. Preselection reduced variables to fewer than 150.

Model training and validation involved nested cross-validation on ProstateNET and external validation on ProstateX2. ML classifiers included Support Vector Machine (SVM), Random Forest (RF), L1-lasso, and Boosted Generalized Linear Model (GLM). Metrics such as AUC, F1-score, precision, sensitivity, and specificity evaluated performance. Of 480 models, the best in nested cross-validation combined Boruta with Boosted GLM (AUC = 0.71, F1 = 0.76). In external validation, L1-lasso with Boosted GLM performed best (AUC = 0.71, F1 = 0.47).

ADC-derived features exhibited the highest discriminatory power, while T2w-derived features were less informative. Combining features from both sources did not enhance performance. The top-performing feature selection methods were Boruta, RFE, L1-lasso, and Random Forest, with minimal differences in results across ML classifiers.

11.1.3 Knowledge-driven models in health and care

Mechanistic models concern a more classical approach in engineering, biological, physical or chemical sciences. Not surprising, most of the models addressing health-related questions in the last decades have been **based on conceptual analogies** to problems solved in these classical quantitative disciplines. Formulating body functions in terms of mechanisms means spelling out the human **body physiology and pathophysiology** at each level and scale, including all variability in their ‘natural language’ of molecules, cells, sub-organ units, whole organs and the whole body. These mechanisms can then be brought together in a multiscale, multilevel human DT of health and disease. This has the advantage that, if the mechanistic DT is accurate, it will be able to correctly predict the outcome of therapeutic interventions, for which no or only very little data is available, i.e., the model generalizes and

¹⁴⁴ Rieke N *et al.*, NPJ Digit. Med. 2020; 3:119. <https://doi.org/10.1038/s41746-020-00323-1>.

extrapolates faithfully. On the other hand, to arrive at a credible DT, accurate data on all model parameters and their distributions, are required, as well as iterations between forming the DT and validation experiments to evaluate intermediate DT stages. Over the last decades, experimental modalities have been considerably advanced so that accurate data acquisition is now possible for many aspects (*e.g.*, imaging, omics, *cf.* previous chapter). **Knowledge-driven DT development** and **experimental modality development** as well as **experimental design** go hand in hand, guiding and promoting R&D in each other. Experience accumulated over decades in responding to biological and clinical questions has led to a large body of DTs and has advanced our understanding on how such twins should be built. Within a network of collaborations between experimentalists, clinicians, and DT developers, and augmented by AI with knowledge-driven DT components, integrated multiscale and multilevel DTs are within reach.

Box 15: Success story – using Digital Twins to advance mechanistic insights.

Success story: Digital Twins for planning partial liver resection during tumour treatment

Status: TRL 3-4

Website: <https://team.inria.fr/simbiotx/software-2/computix/>

Digital Twins can be used to identify discrepancies between mechanistic ideas underlying disease processes and data, which in turn can lead to corrections of these mechanistic ideas with potential consequences on diagnosis, therapy or prognosis. A successful example is ammonia detoxification after drug-induced liver injury (DILI) by overdosing paracetamol (acetaminophen, APAP). Hyperammonemia, a too high ammonia blood concentration can lead to encephalopathy and irreversible brain damage. A DT implementing the consensus set of mechanisms¹⁴⁵ was not able to quantitatively reproduce experimental measurements in an animal model and in a perfused extracted liver¹⁴⁶. The discrepancy could be removed within the DT by adding a hypothetical ammonia sink mechanism. A concrete candidate mechanism (replacement of the irreversible by a reversible Glutamate dehydrogenase (GDH) reaction) could be identified, and by a correspondingly modified DT be shown to quantitatively explain the data¹⁴⁷.

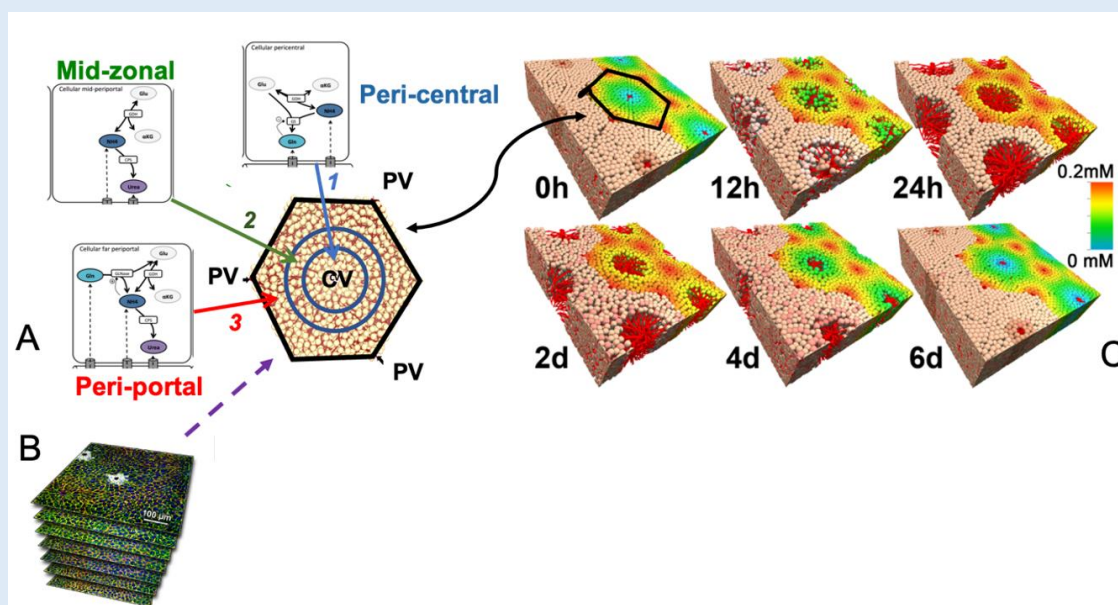


Figure 7: A virtual liver sample twin. (A) Liver is composed of thousands (mouse) to millions (human) repetitive anatomical and functional units, called lobules, that can statistically be approximated by hexagons. Within each lobule, metabolism of ammonia is zoned i.e., different reactions are executed in the peri-portal, mid-zonal and peri-central region (indicated by the italic numbers). The precise lobular micro-architecture has been computed from confocal micrographs (B) A set of reactions in each zone of a tissue sample (C) can be solved to simulate the detoxification of blood from ammonia (colours).

¹⁴⁵ <https://doi.org/10.1002/j.1460-2075.1983.tb01464.x> and <https://doi.org/10.1111/j.1432-1033.1983.tb07458.x>

¹⁴⁶ <https://doi.org/10.1002/hep.27136>

¹⁴⁷ <https://doi.org/10.1016/j.jhep.2015.11.018>

This mechanism was then experimentally validated, and could be demonstrated in an animal model to provide a potential therapy option if it was triggered in the blood. The DTs used a model integrating a compartment model with a spatial temporal model resolving tissue microarchitecture. The compartment model approximates the liver as a well-stirred reactor. Refined DTs now represent 3D tissue microarchitecture down to subcellular level, and the crosstalk of multiple cell types, and permitting to mimic tissue remodelling in acute damage, disease progression and regeneration including the molecular mechanisms in place¹⁴⁸. In these cases, representation of the spatial organization of cells and vessels inside 3D architecture within a DT is indispensable to arrive at correct outcomes. Models along the same line have been developed for regeneration after partial hepatectomy, tumor initialization, clonal expansion, and fibrotic scar development¹⁴⁹. The simulations at liver architecture have been performed with the software TiSim that is currently at TRL 4¹⁵⁰.

11.2 Models as data transformation services

We have defined a Digital Twin in Healthcare (DT) as a computer model that predicts quantities of interest (QoI) necessary to support decision-making within a Context of Use (CoU) in healthcare (*cfr.* PART 1). These DTs can be generic, population-specific or even fully personalised subject-specific. They could *e.g.*, model a cerebral aneurysm and the QoI could be the risk of rupture of that aneurysm in the next year. The CoU could be fully personalised, or for individuals in a sufficiently stratified population (*e.g.*, caucasian females over 45 years of age without other conditions), or much broader in a large population of all individuals that present themselves with a diagnosed cerebral aneurysm.

Let us now introduce a few **generic concepts** that will help in sharpening the role of models, data, and models as data transformation services. It starts with the notion of a *system* that we study and that we may want to change. The system could be the example of a cerebral aneurysm as a pathology that can present itself in humans. The intended CoU will further define the system, it could be the overall pathologic condition of a cerebral aneurysm in a full population, or the specific aneurysm of a specific individual as presented at a specific point in time.

We now define a **system S as a potential source of data**. To learn about the system, we experiment with it. For instance, with medical imaging we can observe the cerebral aneurysm and extract data from it (location, size, shape). We define an experiment E as the process of extracting data from a system by exerting it through its inputs. With this we can now define a model M for a system S as anything to which the same experiment E can be applied to answer questions about S . Going back to the example of the cerebral aneurysm, an experiment that could be done, but not in the real person, is to see under what pressure loadings the aneurysm would rupture, and then asking what the probability would be for such loading to appear in the next year. Such an experiment could be performed on the model M , noted $E(M)$, and the expectation is that the resulting data of $E(M)$ is close enough to the resulting data of E applied to S , noted $E(S)$. More precisely, the output data error $E(M) - E(S)$ (predicted minus real) should be sufficiently small and clinically acceptable in the intended CoU. When this is the case, we say that the model M of the system S is valid in the intended CoU – where we imply that a clear definition of the experiment E is part of the CoU.

A DT is a mathematical model (whether data-driven or knowledge-driven), which, in the vast majority of cases, is not amenable to analytic solutions. So, performing an experiment on a DT will require running a **simulation on a computer** (or a series of simulations in a computer cluster system). This notion is captured by defining a DT as a computer model, so a mathematical model that will be solved on a computer to experiment with it. An important distinction between the model description (in terms of the mathematical formulation and its coding as a computer program) and the experiment description (so, setting up the specific inputs to the computer program that encodes the model) should be underlined here. When doing this, care should be taken to apply an experiment to a computer model for which it is valid (*e.g.*, by setting input parameters within the range for which the model was validated).

¹⁴⁸ <https://doi.org/10.3389/fbioe.2023.1049564> and <https://doi.org/10.1016/j.isci.2023.108077>

¹⁴⁹ <https://doi.org/10.1016/j.isci.2022.105714>, <https://doi.org/10.1007/s11538-017-0375-1> and <https://doi.org/10.1016/j.jhep.2023.05.016>

¹⁵⁰ <https://hal.science/hal-04211418v1>

Establishing the **validity of a model**, *i.e.*, determining if $E(M)$ is close enough to $E(S)$ within the CoU for the specific QoI, is far from trivial. Remember that a model is always related to the tuple system and experiment. Model validation *always* relates to an experiment to be performed on a system ($E(S)$), which should be very clearly defined by the QoI and the CoU. Demonstrating validity is a key element in the credibility assessment of DTs (*cfr.* PART 4). As noted in the first chapter of this PART 2, ASME VV-40:2018 is viewed as a standard for emerging regulatory pathways in Europe in relation to assessing credibility of DTs.

A final relevant notion is that a **model can also be qualified as a system**. So, a model of system $M(S)$ is by definition a source of data and therefore also a new system S' , *which can be defined as* $S' = M(S)$, and we could create a new model $M'(S')$. For instance, a very complex and computationally demanding multiscale DT could be considered as a system, for which a computationally much cheaper data-driven **surrogate model** could be constructed. Note that we now create hierarchies of models and related experiments, so we must be very careful in keeping track of the original system, *e.g.*, a specific person or a stratified population with a certain pathological condition, for which a model of a model is supposed to produce data that resembles the data for the original system. Clearly, well-defined protocols are needed to keep track of such hierarchies of models.

From the infrastructure point of view, a computer model is then viewed as a **data transformation service**, and a DT is then composed of one or more of those services. In the figure below, the basic idea is presented. Input data would be a set of model parameters, a set of boundary conditions and a set of initial conditions. The model M_{CoU} , valid in a specific CoU, then transforms the input dataset $X = \{\alpha, \beta, \gamma\}$ (defining the experiment E to be performed on the model) to the model output $Y = M_{CoU}(X)$, from which in turn the QoI is extracted, $QoI = f(Y)$.

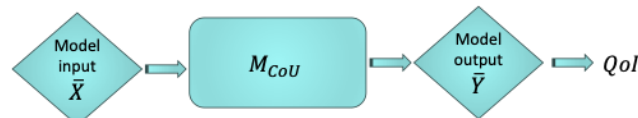


Figure 8: a computer model as data transformation service.

In its simplest form a DT would be like the figure above. In reality, a DT is usually much more complex, where the model M_{CoU} could be executed many times (*e.g.*, for uncertainty quantification), where other data transformation services are needed to prepare the model input (*e.g.*, to extract a geometry of a cerebral aneurysm from an MRI image), where different models are chained together, and where additional data transformation is required to get the final QoI from the model output, and any combination of those.

11.3 Models as data generation services

As noted above, a **model can also be seen as a system itself**, and as such becomes a source of data that can subsequently be used for designing and building other models. For instance, in the example of the aneurysm, one could aim at creating some reduced order model for the pulsatile flow in the aneurysm, and use the output from the original model that computes these flow fields to obtain *e.g.*, a proper orthogonal decomposition on which the reduced order model could be constructed. In this scenario, the model is used as a **data generation service**, and the data is then reused to construct or train other models, possibly relying on AI/ML algorithms. The most obvious reason to do so is when the original model requires huge computational resources, maybe millions of core hours on dedicated tier0 or tier1 HPC systems, and a computationally cheaper surrogate model is needed when fast response is required (*e.g.*, in decision support scenarios), when many instantiations of the model are needed (*e.g.*, for uncertainty quantification), or in *e.g.*, computer aided design scenarios. Many high fidelity time dependent three dimensional multi-{scale, component, organ} models are compute intensive, and to put them to use for *e.g.*, clinical decision support, treatment optimisation, or computer aided design of implants, requires creating (machine learned) surrogate models, and the VHT infrastructure should provision ways to do that.

As noted above it is very important to **keep track of the CoUs/QoI** of the original high-fidelity model. The vision is that output of each run of such high-fidelity simulation is stored in the VHT infrastructure, ready to be reused, possibly by independent third parties and for their own purposes. In such a scenario the third party may not even be aware of the origin of the simulated dataset (the model that created the data) and could use that data as a ground truth for training surrogates. When the CoUs of the simulated data is not kept clearly connected to that of the original model, through metadata, there is the danger of using simulated data beyond its CoU.

11.4 Models as data flow orchestrations

We have two basic objects, a set of models M and a set of data objects D . They are connected as a directed bipartite graph, meaning that elements of M (so, individual models) take as input elements of D and produce as output elements of D that in turn can serve again as input to other elements of M . Consider e.g., a simple workflow, with 2 models $M = \{m_1, m_2\}$ and five datasets $D = \{d_1, \dots, d_5\}$ chained together as shown in the figure below. One can also view the resulting graph as a coloured graph, with data in orange and models in blue, and each edge in the graph going from colour to another.

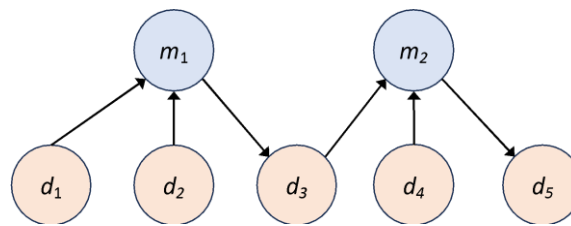


Figure 9: Example of workflow graph

If we now assume that the M and D form the full set of model and data objects in the VHT, and the set of edges E describing all input-output relations between data and models, we can denote the graph $G_{VHT} = (M, D, E)$ as the formal **representation of the collected information available in the VHT**. G_{VHT} is a dynamically growing object, as we keep adding data and models to the VHT and link them up to the G_{VHT} .

We should **not confuse the (ever growing) G_{VHT} with the final model of the full pathophysiology** of a human being. G_{VHT} may contain many different models that compute more or less the same QoI but in different CoUs, or at different spatio-temporal scales, or at different clustering, or operating at different levels of credibility. However, G_{VHT} does contain many different DTs, expressed as workflows. In the example of the figure above, we could imagine that this specific small graph is embedded in the G_{VHT} , and view this as a workflow with three input data objects $\{d_1, d_2, d_4\}$, one intermediate dataset $\{d_3\}$, one output data set $\{d_5\}$ and two models $\{m_1, m_2\}$. We could write $G_{DT} = (M_{DT}, D_{DT}, E_{DT})$ and $G_{DT} \subset G_{VHT}$.

So far, G_{VHT} is a growing but otherwise static object. However, models execute, produce new data, and if new data is available, other models could execute, as described in the first chapter of this PART. The idea would be that the models in the VHT, as it were, scavenge on data, and once new data becomes available for them, they execute (so, a dataflow execution model). To add such **dynamic (eager) model execution** to the whole picture we can resort to a Petri Net, which is “a directed bipartite graph that has two types of elements: places and transitions. A place can contain any number of tokens. A transition is enabled if all places connected to it as inputs contain at least one token”, and “a transition of a Petri net may fire if it is enabled, i.e., there are sufficient tokens in all of its input places; when the transition fires, it consumes the required input tokens, and creates tokens in its output places.”¹⁵¹ In our case, places are data objects, tokens are new data items arriving or computing and storage resources becoming available, and transitions are the models.

¹⁵¹ https://en.wikipedia.org/wiki/Petri_net

12 Integration of resources

12.1 Identification of possibilities for integration

The main resources, data and models, can be integrated in a number of ways, and can then be mapped to the computing, storage, and networking resources. Data can be grouped together, pooled, to form richer datasets, maybe forming a new DOT. Models can be integrated into multiscale/multiorgan models. Integrating chains of models and data together where the output data of one model serves as the input data to another model, results in workflows, which are considered the third basic resource in the VHT (next to data and models). Likewise, computational and storage resources can (dynamically) be pooled together to provide the required infrastructure to interrogate data, store (new) data, and to execute workflows.

12.1.1 Integration of multiscale models

Multiscale models encompass model features at multiple scales across space and/or time. The Multiscale Modelling and Simulation Framework (MMSF) was defined^{152,153} based on the Multiscale Modelling and Simulation Language (MMSL)¹⁵⁴, which describes the coupling of many sub-models to the multiscale model. There is an increasing **integration of AI/ML and multiscale modelling**¹⁵⁵, where machine learning explores big parameter spaces issued from complex heterogeneous data for the identification of correlations, and multiscale modelling predicts the system dynamics and identifies causality. In such combined models, system validation is of prime importance to avoid that the AI/ML algorithms lead to non-physical or non-biological solutions.

Prospectively, markup languages, permitting a unique description of models, should be able to describe all knowledge-driven models, including even agent-based models of individual cells that are becoming increasingly sophisticated. MultiCellDS or MultiCellML may be promising steps into this direction.

Tightly coupled multiscale models, where models interact in cycles (*e.g.*, a microscale model providing input data to a macroscale model, that in turn at a next time step provides input data to the microscale model, etc.) require dedicated coupling frameworks, as discussed above, and from the point of view of the VHT are considered a new model. In contrast, **loosely coupled multiscale models** do not have such cyclic feedback loops and can be composed as a workflow, and are therefore viewed as a workflow resource.

Regardless of the chosen integration method, attention needs to be paid to the **build-up of errors** coming from the combination of models (even if the individual models were perfect). This area is understudied in general and fundamentally important for the VHT.

Box 16: Success story – Virtual Twins for Cancer.

Success story: Virtual Twins for Cancer

Status: *research phase (TRL 4)*

Website: <https://permedcoe.eu/use-cases/>

Virtual Twins for cancer connects molecular, cellular, and tissue scales to aid in patient-specific prognosis, enhance early detection, and define treatments aimed at improving survival rates¹⁵⁶. The use of omics data — including genomics, transcriptomics, and proteomics — is central to the personalisation of these models. Incorporating comprehensive molecular data from individual patients allows the Virtual Twins to offer a more accurate representation of tumour behaviour and response to treatments. These models can identify biomarkers and points of intervention that are specific to each patient and rank potential drug activities to provide personalised treatments. This approach enables a deeper understanding of the disease mechanism and facilitates the development of targeted and effective therapies. Continuous integration and analysis of omics data ensure that the models remain up-to-date and reflective of the latest scientific insights, further enhancing their predictive power.

¹⁵² Borgdorff *et al.*, Proc.Comput. Sci.2013 ;18,1097–1105. <https://www.doi.org/10.1016/j.procs.2013.05.275>

¹⁵³ Chopard *et al.*, Philos Trans A Math Phys Eng Sci. 2014; 372(2021): 20130378. <https://www.doi.org/10.1098/rsta.2013.0378>

¹⁵⁴ Veen & Hoekstra, Proceedings of ICCS, 2020; 425-438. <https://hdl.handle.net/11245.1/298a965c-72ae-4771-9069-d4a6b2b8ea26>

¹⁵⁵ Alber *et al.*, npj Digit. Med. 2019;2, 115. <https://doi.org/10.1038/s41746-019-0193-y>

¹⁵⁶ <https://www.nature.com/articles/s41540-023-00314-4>

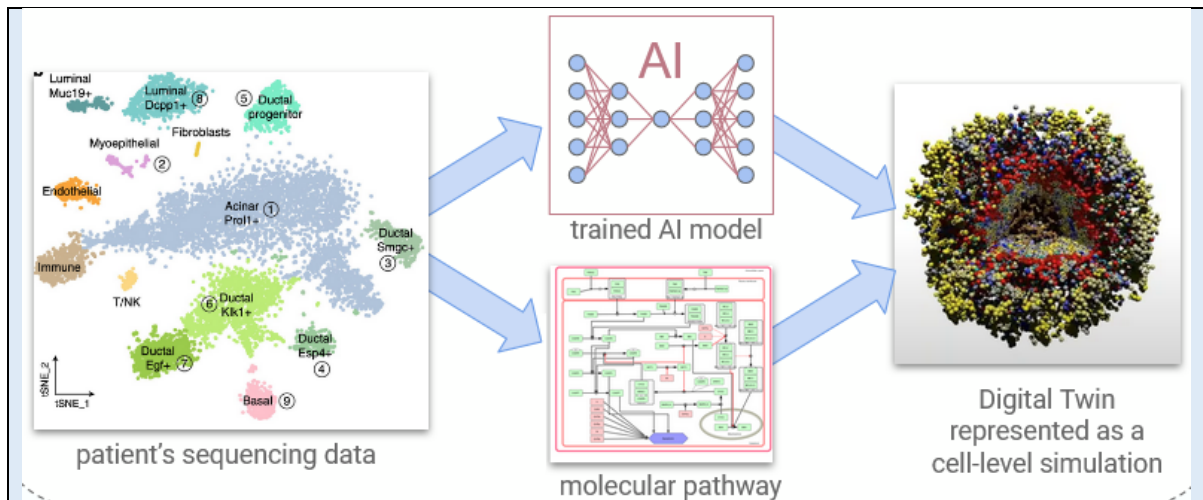


Figure 10: Integration of patient's sequencing data and molecular mechanisms in a cell-based simulation framework to capture individuals' genetic and environmental influences in order to provide a deeper understanding of the biological context and response to therapies.

This effort requires collaboration from researchers across multiple disciplines, including biochemistry, biology, physics, and computer science, as well as clinicians. Collaboration between researchers and clinicians is fundamental for both the development and practical implementation of these models. This synergy ensures that the models are scientifically robust, clinically relevant, and user-friendly.

Additionally, the Virtual Twins for cancers need enabling infrastructures for patient data gathering and analysis, such as those provided by the European Health Data Space (EHDS) and the European Open Science Cloud for Cancer (EOSC4Cancer). High-performance computing (HPC) centres like Marenostrum5, part of the EuroHPC pre-exascale systems, are also essential for the complex simulations required.

Box 17: Success story – a multi-level atherosclerotic plaque growth model for coronary arteries.

Success story: a multi-level atherosclerotic plaque growth model for coronary arteries

Status: TRL 5

Website: <https://pubmed.ncbi.nlm.nih.gov/31946985/>

This computational model predicts atherosclerotic plaque growth in coronary arteries by simulating plaque progression and arterial wall deformation. It integrates clinical, biochemical, and imaging data, including coronary computed tomography angiography (CCTA) imaging and blood biomarker analysis at baseline and follow-up.

The model incorporates multiple components: (i) **Blood Flow and Plasma Modelling:** Blood flow is modeled using the Navier-Stokes equations, while plasma flow in the arterial wall is simulated as a porous medium via Darcy's law. (ii) **Lipoprotein and Monocyte Transport:** Transport of LDL and HDL is modeled using the Kedem-Katchalsky equations, with monocyte infiltration driven by cytokines. (iii) **Atherosclerotic Processes:** LDL oxidation, foam cell formation, smooth muscle cell (SMC) proliferation, and cytokine-mediated SMC differentiation and collagen production are included to simulate plaque components. (iv) **Arterial Wall Thickening:** Stress-strain equations simulate wall deformation and thickening caused by plaque growth.

Model implementation involved: (i) **Data Collection:** Clinical data, blood biomarkers (e.g., LDL, HDL), and CCTA imaging at two time points. (ii) **3D Reconstruction:** Arteries were segmented using a level-set algorithm, with geometries registered between baseline and follow-up. (iii) **Simulation:** Plaque growth simulations were performed on baseline geometries, producing deformed arterial walls and lumen areas. (iv)

Simulated geometries were validated against real follow-up imaging using statistical analyses (e.g., regression, Bland-Altman plots), achieving a correlation ($r^2 = 0.49$, $P < 0.0001$) between simulated and actual geometries. The model effectively identifies regions prone to plaque progression and predicts growth rates, demonstrating potential for assessing high-risk coronary artery disease patients and enhancing clinical decision-making.

12.1.2 Workflows

Workflows allow for the fast deployment of the VHT in different computing architectures (such as Cloud and HPC, *cfr* PART 3) and enable the reproducibility of use cases across different users.

Technically, the use of workflows greatly **facilitates the deployment** of complex and large use cases **in different computing facilities** and the **generation of intermediate data structures** in case the simulation has execution errors or perturbations (as energy outages). They also enable **repeatability and reproducibility** of use cases while providing a way to integrate provenance of data and privacy by allowing for federated workflows across multiple sites. In the VHT, automatically keeping track of provenance, certainly in relation to the credibility of the output data of a workflow, is a key requirement. Workflows **facilitate user understanding** of the overall DT without the need of delving into detailed model documentation and metadata, hence lowering the learning curve to be able to simulate use cases by themselves.

For example, a workflow recipe for a complex use case for modelling brain tumours combining different types of models, data sources, and output data for a commonly used workflow manager can allow researchers to seamlessly deploy exactly the same use case in the Cloud (*e.g.*, AWS and Google Cloud) or in different high-end HPC clusters (*e.g.*, LUMI, MareNostrum) with minimal amount of work.

Even more so, the use of workflows usually but not necessarily involves **the division of the work in containerised pieces** (such as Docker, Apptainer (former Singularity)). This compartmentalization allows for the decoupling of the complexity of each part of the workflow and its usability in complex use cases. Likewise, these containers facilitate benchmarks of different tools aimed at the same or similar tasks by solving unit tests to study *e.g.*, the validation, verification and uncertainty quantification (VVUQ) of each model and implementation and also allow for the definition of federated workflows by clearly dividing the work of each software piece and assigning them a given computing cluster (one part in the cloud, another on-site *etc.*). Also automating this process, or creating a demand/supply market for semi-automated execution of complex DTs, in relation to the dataflow mechanisms introduced earlier, can be facilitated by advanced VHT workflow engines.

Despite this, the organisation of use cases in containers and workflows is a **clear overhead on the researchers' side** and currently there are very few ways that this additional effort can be leveraged. There are a myriad of different **workflow managers** in the biomedical domain with slightly different focuses; some of them are open source, while others are commercial. For these reasons, the community has been working on delivering **standardised workflow descriptions** to facilitate the deployment of workflows' recipes across managers. Likewise, workflows registries have been established to facilitate the discovery and re-use of workflows in an interoperable way, as well as frameworks providing tools for construction and execution of workflows and backend programs for the reproducible execution of standard workflows. These elements are discussed in more detail in PART 4.

12.1.3 AI tools for resource integration

The integration of AI tools presents significant opportunities for **resource orchestration and optimization** within the VHT. AI can assist researchers in selecting appropriate building blocks for assembling a DT (similar to recommendation systems used in other domains), optimizing computational workflows, harmonizing data formats, supporting standardization efforts and enabling seamless integration of diverse data and models. By leveraging a machine-readable overview of available resources, pricing, and computational requirements, AI tools could optimize the execution of VHT computations in terms of cost and speed, enhancing efficiency and utility.

Regarding **data-model integration**, standardized systems for cataloguing data and models could enable AI to support the identification of suitable input data for VHT applications. Missing data could potentially be generated through deterministic or AI-based methods. Similarly, model compatibility could be assessed by defining input-output specifications for DTs within the VHT. AI could further enhance this process by retrieving and curating relevant literature and data, ensuring that the outputs of one model align with the inputs of another.

12.2 Integration of resources as a basis for knowledge finding

Another example of integration of resources is the exploration of the VHT as a knowledge base, using NLP in general and, given their recent advances, large language models (LLMs) in particular. With the

ever-increasing volume of resources in the VHT, combined with the research, legal, ethical, privacy-related, and regulatory documents in this field, there is a growing need for efficient tools to extract knowledge and insights from this wealth of information. This **knowledge retrieval** can take the form of a **knowledge graph** (KG) or model. These knowledge graphs might provide feedback to the VHT platform users on *e.g.*, resources that can be coupled or specific data needs that would allow to generate a more comprehensive DT. Additionally, NLP can produce narratives of the platform outcomes by automatically generating reports that contain the required technical or clinical information, tailored for the targeted user, *e.g.*, the clinician, or even high-level stories understandable by the patient. Integrating frameworks of Large Language Models (LLMs) and Knowledge Graphs (KGs) is key to enable seamless exploration and exploitation of the VHT platform for diverse applications.

12.2.1 Large Language Models

With the ever-increasing volume of resources in the VHT (models, datasets, workflows, etc.), combined with the research, legal, ethical, privacy-related, and regulatory documents in this field, there is a growing need for efficient tools to extract knowledge and insights from this wealth of information. To this end, LLMs can provide powerful capabilities for knowledge retrieval, summarization, and user guidance. Their ability to process unstructured text and generate context-aware narratives can enhance the accessibility and usability of the VHT. The following applications of LLMs can be imagined.

1. **Knowledge retrieval from VHT knowledge base:** LLMs can retrieve, process, and consolidate relevant information from the VHT's knowledge base – an extensive repository of scientific, legal, ethical, and regulatory documents – acting as a Retrieval-Augmented Generation (RAG) tool. This capability can allow users to obtain answers to queries posed in natural language queries, for example: *"What are the regulatory requirements for using AI-based models in healthcare?"*, and *"What computational modelling methods have been applied to atrial fibrillation?"*.
2. **Summarization of complex content:** LLMs can generate concise summaries of detailed technical, regulatory, or clinical content tailored to the user's needs. For example, a researcher may request *"a summary of recent advancements in cardiovascular Digital Twin modelling"*, or ask about *"a summary of regulations that apply to in-vitro medical devices"*.
3. **Automated report generation:** LLMs can compile structured reports summarizing the outputs of simulations, experimental findings, or compliance checks. These reports can cater to diverse stakeholders: (i) **clinicians**: simplified, actionable insights for patient care; (ii) **researchers**: detailed descriptions of model parameters and assumptions; (iii) **regulatory bodies**: traceable evidence of compliance with standards and guidelines; (iv) **patients**: simplified high-level stories to explain progress of the disease or its treatment.
4. **Guidance and contextual support:** LLMs can assist the different users in navigating the VHT platform, providing a range of functionalities, depending on the users' needs. A **helpdesk** functionality tailored towards model developers, data providers and VHT composers can help them with tasks such as uploading and combining resources. An **educator** functionality tailored towards citizens, medical professionals and other stakeholders can provide training and explanations, clarifying terminology, or suggesting relevant documents and resources for further exploration. An **accessibility** functionality can serve users of different languages, but also translate electronic health records, which are usually maintained in national/regional languages.

By enabling robust retrieval and summarization capabilities, LLMs can make the VHT platform more accessible to users with varying levels of expertise. Although the above LLM applications can be developed by directly applying LLMs to the VHT knowledge base, their robustness can further be enhanced by integrating knowledge graphs into the retrieval process for handling complex queries.

12.2.2 Knowledge graphs and LLMs: A Synergistic Framework

Knowledge graphs, which are structured representations of information capturing relationships between entities in a graph format, can serve as a semantic layer connecting the platform's core resources – data objects, model objects, workflows, and standards – to user-defined queries. KGs organise data into nodes (entities) and edges (relationships) with associated semantics, thus enabling efficient querying and reasoning over complex datasets. By combining the structured reasoning capabilities of KGs with the interactive potential of LLMs, users can intuitively access, simulate, and analyse complex scenarios. For example, the VHT user may pose an instruction of *"Test a new transcatheter aortic valve design on elderly women with diabetes."*, and the KG/LLM framework generates an answer identifying all the relevant resources/assets and instructions on how to achieve the desired testing through the available infrastructure.

Acting as an integrated representation of all key resources in the VHT ecosystem to enable advanced semantic querying and reasoning, the **construction of KGs** can leverage the following VHT features:

- **Data objects**, data sets annotated with metadata based on FAIR principles, capturing demographic, clinical, physiological, and imaging data as nodes in the KG. They use proposed ontologies to define relationships such as “used in”, “validated by” or “derived from”.
- **Model objects**, as previously discussed, define relationships such as “input to”, “output from” or “compatible with” data objects and workflows, as well as providing a link to validation and performance metrics.
- **Workflow objects**, being orchestrated sequences of actions combining data and models for simulations and defining relationships such as “requires”, “produces” or “meets regulatory standards”.
- **Ontologies and standards**, using existing ontologies for the VHT resources (for example data objects), using standardized vocabularies (e.g., ISO, DICOM) to annotate all entities and relationships and ensuring interoperability and traceability within the KG.

The following applications can be imagined for **KG exploitation**.

User-Centric Query Resolution: by integrating KGs and LLMs, users can gain an intuitive interface for exploiting the VHT platform. The system can translate user queries into actionable workflows, leveraging the KGs to identify and link relevant resources.

1. **Example Query:** *"Test a new transcatheter aortic valve design on elderly women with diabetes."*
2. **Workflow:**
 - a. **Input Parsing:** The LLM parses the query, identifying key entities (e.g., "aortic valve design," "elderly women," "diabetes").
 - b. **Resource Identification:** The KG retrieves: (i) Relevant datasets matching the demographic and clinical attributes; (ii) Compatible models for simulating valve performance under these conditions; (iii) Workflows for combining data and models.
 - c. **Guidance:** The system generates step-by-step instructions, including: (i) Input parameters for the models; (ii) Tools to preprocess data; (iii) Simulation execution instructions; (iv) Methods to analyse and interpret results.
 - d. **Validation:** User compares simulation results against any available experimental data for the new device. The KG ensures that resources and workflows meet relevant regulatory standards.

Enhanced Reporting and Summarization: the KG-LLM integration can also provide tailored summaries and visualizations of the results, ensuring accessibility for diverse stakeholders. Researchers can access detailed technical analyses. Clinicians can receive concise, actionable insights.

Gap Analysis and Resource Optimization: the KG framework can enable systematic identification of missing resources in the VHT platform, fostering continuous improvement.

- **Gap Identification:** Analyse failed or incomplete queries to infer missing datasets, models, or workflows. LLMs can even invent new representative queries to identify gaps. Highlight under-represented demographics or conditions.
- **Resource Prioritization:** Recommend areas for data acquisition or model development. Suggest updates to workflows or standards to enhance coverage.

The KG framework can extend its utility to *other high-impact areas*:

- **Regulatory Support:** traceability of workflows and results to ensure compliance.
- **Collaboration:** foster connections between researchers by linking complementary resources.
- **Knowledge Discovery:** enable hypothesis generation by identifying novel patterns in connected data and models.

The integration of resources into a KG framework within the VHT platform will align with the broader vision of enabling personalized, predictive, and precise healthcare. By leveraging the semantic structure of KGs and the interactive capabilities of LLMs, users will be able to intuitively explore the platform, execute simulations, and generate actionable insights. Furthermore, the KG's ability to identify gaps will ensure that the platform evolves dynamically to meet future healthcare challenges.

Box 18: EDITH development – EDITH-CSA knowledge sourcing and knowledge graphs.

EDITH development: Knowledge sourcing

Status: *deployed, accessible via* <https://www.edith-csa.eu/edith-knowledge-base/>

To showcase the knowledge retrieval and summarization discussed above, an LLM web application¹⁵⁷ has been developed as part of EDITH CSA and has been made publicly available. The application is built on a VHT knowledge base consisting of approximately 3,000 documents, which comprise of approximately 2,700 scientific documents and approximately 300 legal/regulatory documents. Furthermore, to automatically and continuously update the knowledge base, an integration with the OpenAIRE VHT community¹⁵⁸, which automatically searches for publications relevant to the VHT initiative, has been made. The application, powered by frontier LLM models, exploits semantic similarity search to fetch relevant documents for responding to user queries. It allows users to (i) ask questions in natural language, and (ii) ask for generation of thematic summaries on user-specified themes. The responses are always drawn exclusively from the literature included in the knowledge base—thus minimising hallucinations—and references to the specific document are included in the response.

On top of the same OpenAIRE knowledge base, a first proof of concept towards a knowledge graph was made, albeit restricted to the analysis of scientific literature metadata. In this pilot implementation, a KG of approximately 40K nodes (representing publications, authors, author affiliations, keywords, and fields of studies) and approximately 60K relationships (for example a 'publication node'-[:has author (relationship)]-> 'author node') was created. This KG was then exploited for advanced search and filtering, data visualisation, keyword/author specific analysis, collaboration insights, author and publications rankings, research trend analysis (for example, uncovering trends in keywords for emerging areas), identifying interdisciplinary connections and collaboration networks, identifying paths connecting authors (to identify common grounds and collaboration opportunities), community detection, and clustering.

¹⁵⁷ <https://www.edith-csa.eu/edith-knowledge-base/>

¹⁵⁸ <https://dth.OpenAIRE.eu/>

13 Digital Twin use cases

Box 19: Use Case - Glycemic control in ICU patients (Digital Twin)

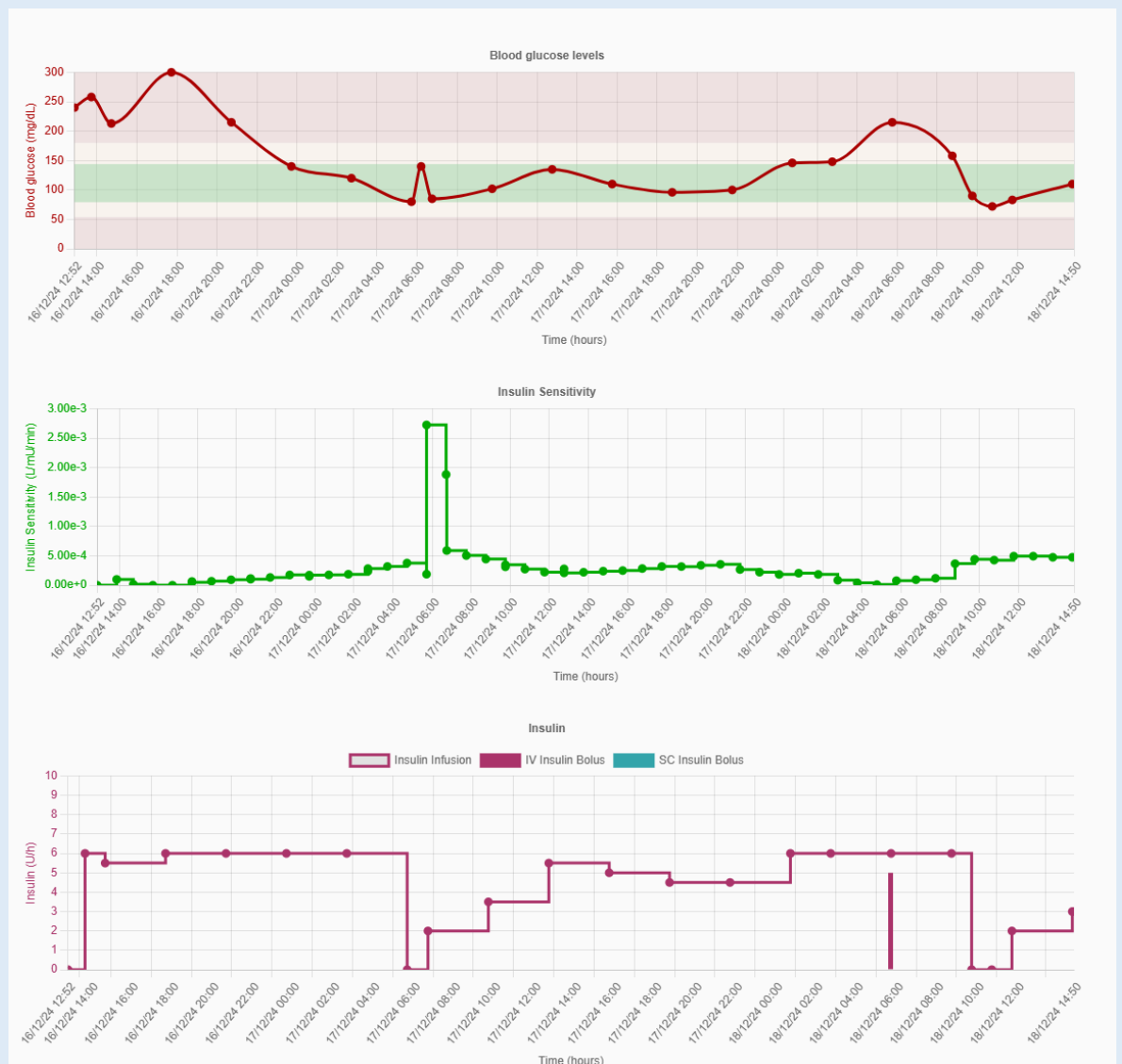
Use Case: *Glycemic control in ICU patients*

Status: *start-up company, solution in clinical trial*

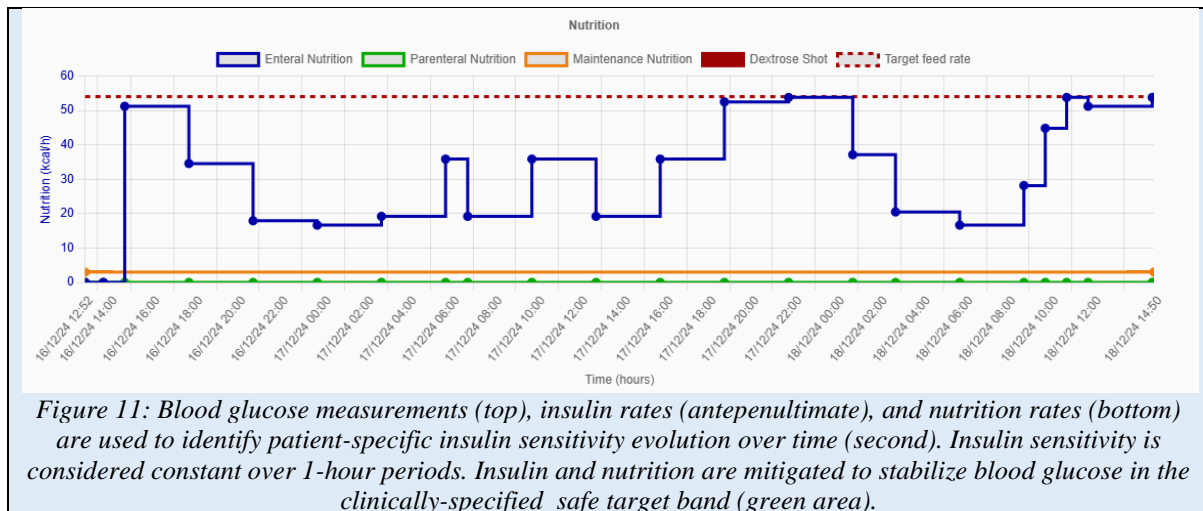
Website: www.insilicare.com

InSiliCare's AI-powered Digital Twin is a clinical decision support solution for ICU patients requiring glycemic control. It provides personalized, safe, and effective management of blood glucose levels and nutrition delivery. Insulin and nutrition treatments are calculated to maximise safety from hypoglycemia, while controlling patient blood glucose levels and optimizing nutrition towards a configurable physician-determined practice of care.

The InSiliCare Digital Twin is a clinically validated mathematical model characterizing the insulin-glucose pharmacodynamics and pharmacokinetics and estimating patient-specific, time-varying **insulin sensitivity** in adult ICU patients¹⁵⁹. Insulin sensitivity is a time-varying parameter characterizing patient-specific metabolic response to insulin-mediated glucose uptake. This “white-box” model consists of 5 primary ordinary differential equations and can be coupled with predictive-AI models to simulate patients' future response to treatment. Model's inputs to identify insulin sensitivity are glucose assays, insulin and nutrition rates, as well as patient diabetes status.



¹⁵⁹ <https://doi.org/10.1016/j.cmpb.2010.12.008>

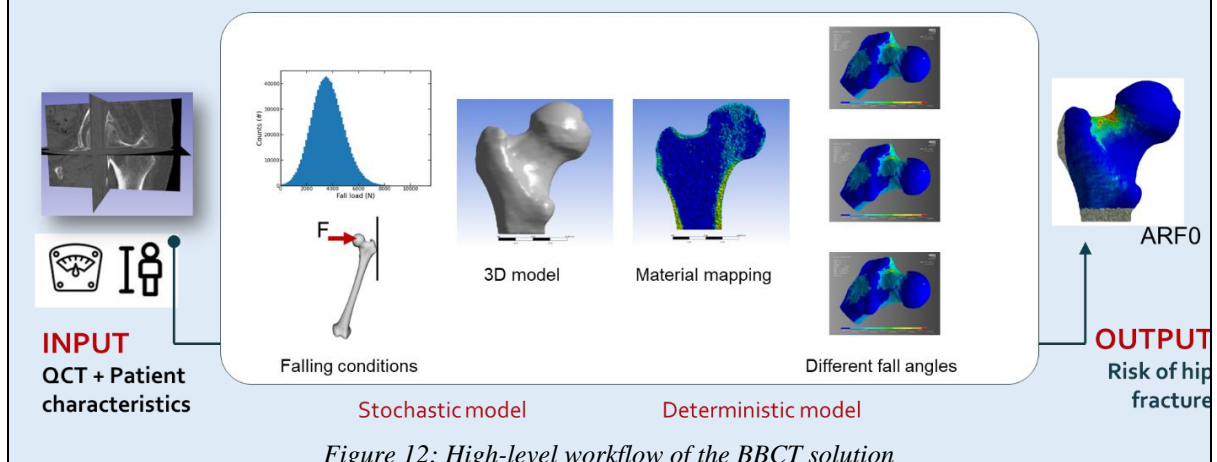


Box 20: Use Case - Osteoporotic fracture risk prediction (Digital Twin)

Use Case: Bologna Biomechanical Computed Tomography for osteoporotic fracture risk prediction
Status: clinical research studies

Bologna Biomechanical Computed Tomography (BBCT) is a Digital Twin methodology designed to predict the mechanical strength of the femur under critical loading conditions in osteoporotic patients. Quantitative Computed Tomography (QCT) scans of the hip region and patient data inform a subject-specific Finite Element (FE) model. One million potential fall scenarios are then simulated based on a body with the height and weight of the patient. The resulting impact forces are estimated. Variables—such as initial and final velocities, accelerations and impact attenuation—are modelled as normal distributions informed by literature data. Subsequently, the patient-specific model is run 28 times, varying the femur's orientation at the impact. To replicate a sideways fall, a concentrated force is applied at the centre of the femoral head, while a contact interaction is defined between the surface of the greater trochanter and a rigid plane oriented perpendicular to the force. The distal part of the femur is constrained to a hinge located at the knee's centre. For each impact orientation, the force intensity required to fracture the femur is computed based on principal strains in a region of interest located proximally. By comparing the FE-derived failure loads to the estimated impact forces, the absolute risk of hip fracture (ARF0) is calculated.

In pre-clinical studies on cadaver preparations the Digital Twin showed an accuracy in predicting the force required to fracture a proximal femur in side-fall conditions of 85%. In a retrospective clinical validation on a cohort of 100 women, half with a hip fracture, ARF0 showed a stratification accuracy of 87%, compared to 75% for the clinical standard of care.



Box 21: Use Case - Universal Immune System Simulator for Tuberculosis (Digital Twin)

Use Case: Universal Immune System Simulator for Tuberculosis

Status: augmenting clinical trials

Website: <https://www.mimesis.srl/>

The Universal Immune System Simulator (UISS) offers a robust, clinically validated platform to address these challenges by simulating and predicting immune responses *in silico*. UISS-TB has demonstrated its utility in optimizing therapeutic strategies for Tuberculosis, evaluating drug regimens, and predicting outcomes.

UISS-TB employs a comprehensive agent-based model that simulates the immune system's interaction with Mtb. The model integrates:

- Cytokine and chemokine signaling pathways.
- Immune cell dynamics (e.g., macrophages, T-cells).
- Bacterial growth and host-pathogen interactions.

This 'white-box' model ensures transparency and reproducibility, leveraging real-world data such as cytokine levels, bacterial load, and patient demographics to deliver accurate predictions. Hosted on secure virtual machines, UISS-TB safeguards data privacy while supporting integration with electronic health records (EHRs) for automated data input.

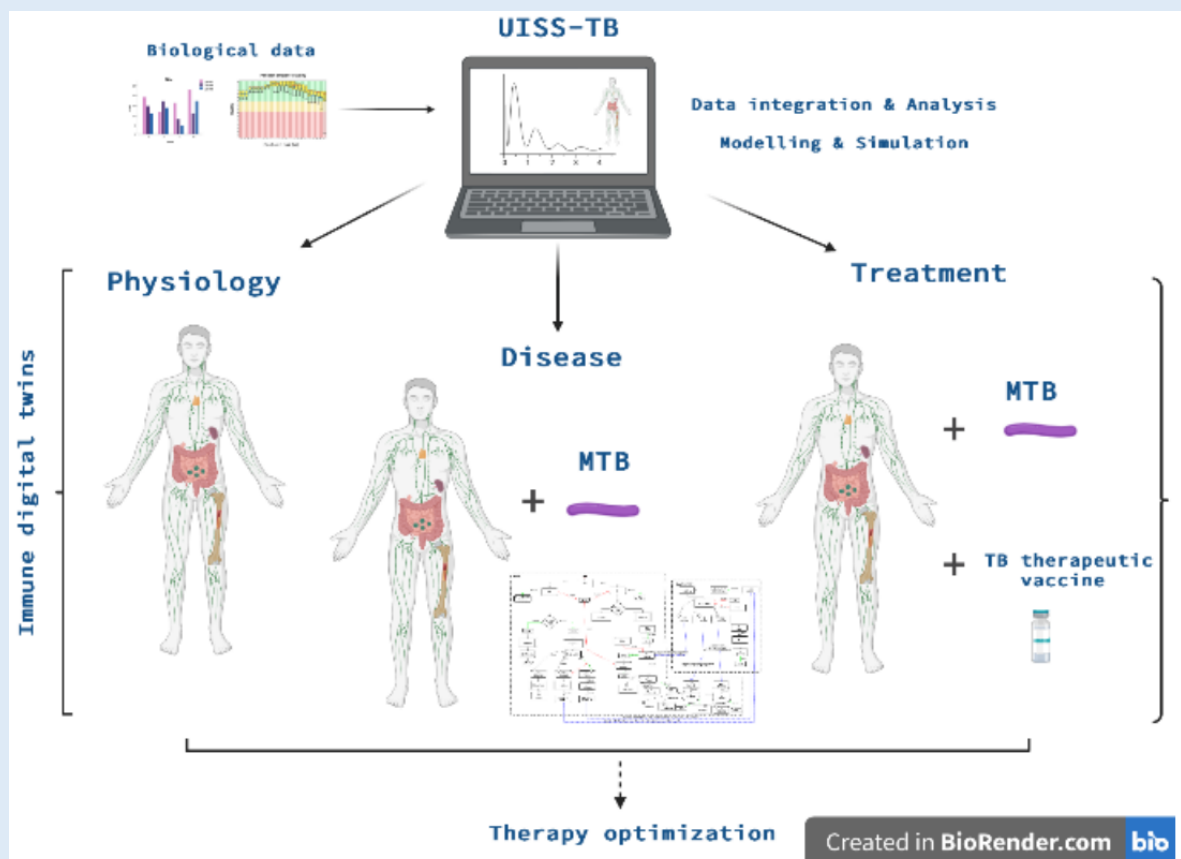


Figure 13: UISS-TB layers and their implementation, aimed to simulate *Mycobacterium Tuberculosis* (MTB) dynamics and its interaction with the human immune system along with the therapeutic interventions.

Box 22: Use Case - Epileptogenic zone localisation for surgical planning in epilepsy patients (Digital Twin).

Use Case: Epileptogenic zone localisation for surgical planning in epilepsy patients

Status: in clinical use

Website: <https://www.cloudsofcare.com/>

The **Digital Twin model in Persyst ESI powered by Epilog** combines **EEG** and **MRI data** using advanced algorithms to create a virtual representation of the patient's brain¹⁶⁰. It helps identify the epileptogenic zone by solving mathematical equations that model the electrical fields in the brain.

The analysis comprised spike detection, clustering and source localization using an individualized head model. EEG source imaging was done with the finite difference method to calculate the forward model. A $1 \times 1 \times 1 \text{ mm}^3$ head model was constructed from the patient-specific MRI by segmenting it into six tissue classes (air, gray, white matter, cerebrospinal fluid, skull, and scalp). The solution points were restricted to the gray matter in a $4 \times 4 \times 4 \text{ mm}^3$ grid. ESI localization was performed at the onset, half-rising, and peak of the spikes using standardized low-resolution electromagnetic tomography (sLORETA). The resulting 3D visualization of seizure activity aids clinicians in making precise, informed decisions for epilepsy surgery.

Box 23: Use Case - the Atrial Modelling Toolkit for cardiovascular Digital Twins (Digital Twin).

Use case: the Atrial Modelling Toolkit for cardiovascular Digital Twins

Status: for research purposes

Website: <https://github.com/pcmlab/atrialmtk>

The Atrial Modelling Toolkit¹⁶¹ (**atrialmtk**) aims to overcome the challenges of constructing cardiac models at scale through the development of a robust, open-source pipeline for bilayer and volumetric meshes for atrial models.

The computer model takes segmentation masks of the atria, which can be produced from raw MRI or CT data by an expert or input directly where the user already has them, or existing atrial surface meshes as input. After the identification of some key anatomical landmarks, the model automatically produces a simulation grade atrial mesh that incorporates atrial regions and atrial fibres, which are input into an electrophysiological simulation.

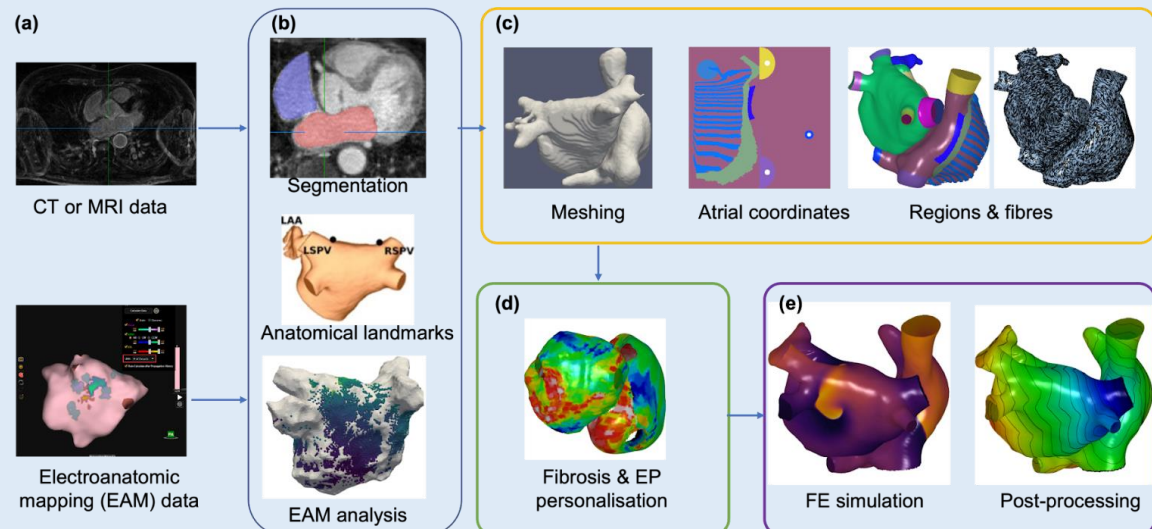


Figure 14: Atrialmtk model pipeline¹⁶²

An atrial fibrillation virtual cohort study was recently performed with 100 patient-specific models constructed from late-gadolinium enhancement magnetic resonance imaging data to demonstrate the potential of combining personalized biophysical simulations with machine learning approaches to predict long-term ablation therapy

¹⁶⁰ <https://doi.org/10.1111/epi.17460>

¹⁶¹ <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2023.0038>

¹⁶² Roney, C.H. et al. (2023), Constructing bilayer and volumetric atrial models at scale. Interface focus, 13(6), p.20230038.

outcome¹⁶³. The Digital Twin model construction pipeline takes imaging data (MRI or CT) together with electroanatomic mapping data to construct anatomically detailed models, calibrate these models, and simulate the response to different treatment approaches¹⁶⁴. The model construction pipeline includes automated segmentation¹⁶⁵, meshing, structural fibre maps from human ex-vivo diffusion tensor MRI¹⁶⁶ and calibration to electrical conduction properties¹⁶⁷.

Machine Learning is used to complement the obtained biophysical simulations in an in-silico trial to predict long-term response to treatment strategies for patients with atrial fibrillation. Atrial fibrillation, characterized by irregular activation of the heart's top two chambers, can be treated with ablation to target tissue sustaining the arrhythmia. However, since only one ablation strategy can be applied to each patient, determining the optimal strategy that treats the arrhythmia while minimizing tissue ablation remains challenging. Our novel computational pipeline successfully predicted long-term atrial fibrillation recurrence in individual patients following ablation therapy by combining outcome data with patient-specific simulation responses. Importantly, classifiers trained on a combination of clinical data and simulation data outperformed those trained on clinical data alone.

¹⁶³ <https://pubmed.ncbi.nlm.nih.gov/35089057/>

¹⁶⁴ <https://pubmed.ncbi.nlm.nih.gov/33041850/>

¹⁶⁵ <https://pubmed.ncbi.nlm.nih.gov/33317334/>

¹⁶⁶ <https://pubmed.ncbi.nlm.nih.gov/32458222/> and <https://zenodo.org/record/3764917>

¹⁶⁷ <https://pubmed.ncbi.nlm.nih.gov/31026761/>

14 Technology for VHT: Conclusions and Recommendations

14.1 Conclusions

Part 2 of the roadmap delves into the technological foundations of the VHT. It explores how to effectively organize and integrate diverse resources to establish a robust and interoperable platform.

The vast array of data and models that constitute the **VHT requires a well-defined organizational framework**. This framework utilizes a multidimensional data space, where each data point is meticulously characterized by a data object profile (DOP) and associated with a specific data object type (DOT). Establishing a consistent and comprehensive schema for the DOT and DOP is crucial to ensure the seamless interrogation, integration, and execution of all VHT resources. This multidimensional approach encompasses six key dimensions: three spatial coordinates, time, clustering, and credibility. The credibility dimension is particularly important, as it provides a quantitative measure of the trustworthiness of the data and models within the VHT. This will play a crucial role in gaining acceptance from healthcare professionals and regulatory bodies.

Four primary types of resources are essential for creating a cohesive and interoperable VHT: **data objects, annotation services, model objects, and workflow objects**. Data objects encapsulate the raw digital data, which are then transformed into meaningful information through annotation services. Model objects, encompassing both knowledge-driven and data-driven (AI) models, play a critical role in generating new data objects from existing ones, thereby driving knowledge expansion. Workflow objects, on the other hand, represent the structured integration of data and model objects, enabling the simulation and prediction of complex biological processes.

Model objects, as a fundamental building block of the VHT, require a standardized approach to ensure their **consistency, reproducibility, and interoperability**. To achieve this, model objects within the VHT are characterized in terms of their system, experiment, and **context of use (CoU)**. Defining the system involves specifying the target biological entity or process being modelled. The experiment outlines the specific conditions and interventions being simulated, providing essential context for interpreting the model's output. Lastly, the CoU clarifies the intended application and limitations of the model, ensuring its appropriate use. The concept of a data-driven modelling can be extended beyond biological entities to encompass models themselves. This enables the creation of surrogate models, which are simplified representations of complex, computationally intensive models, facilitating faster and more efficient simulations.

The **integration of AI** techniques will be instrumental in effectively managing the vast amounts of data and models within the VHT infrastructure. AI-driven solutions can streamline data pre-processing, feature extraction, and quality control, significantly enhancing the efficiency and accuracy of data integration. Machine learning algorithms, particularly deep learning models, can be trained on large datasets to build robust predictive models for disease risk, treatment response, and other clinically relevant outcomes. Natural Language Processing (NLP) techniques can further augment the VHT by facilitating knowledge extraction from unstructured text data, such as scientific literature, regulatory documents, and clinical records. This will enable the automatic retrieval of relevant information and support the creation of a comprehensive and constantly evolving VHT knowledge base.

The **practical applications and potential impact of the VHT are illustrated** through a series of Digital Twin use cases, which are presented throughout the roadmap. These use cases, spanning diverse healthcare domains, highlight the versatility and transformative power of the VHT. Examples include glycemic control in intensive care patients, prediction of osteoporotic fracture risk, simulation of the immune system response to tuberculosis, and personalized treatment planning for epilepsy patients. These use cases demonstrate the potential of the VHT to improve patient care, accelerate drug development, and advance our understanding of human health and disease.

14.2 Recommendations

The successful realization of the VHT hinges not only on a clear vision and a collaborative ecosystem of stakeholders, but also on the **development and comprehensive organization and integration of resources** in a well-defined framework. As such, the following recommendations can be made.

1. **Advancing *in silico* Technologies for a Robust, Versatile and Impactful VHT:** To fully realize the transformative potential of the VHT, substantial investment in the development of *in silico* technologies is crucial. This includes focusing on longitudinal and longer-term human health and disease modelling, capturing the long-term dynamic nature of human physiology. The development of multiscale models is essential for integrating knowledge across different levels of biological organization, from cells to organs to the whole body. Incorporating new data modalities beyond traditional clinical data will enrich the VHT, enhancing its predictive power. Omics integration will provide a comprehensive view of individual biological processes, enabling personalized insights into health and disease. AI will play a central role in the VHT, not only in data analysis and model generation but also in driving the interaction between AI and mechanistic modelling.
2. **Data-Generating Hardware Addressing a Critical Need for VHT:** The development of advanced data-generating hardware technologies is essential for the progress of VHT applications. Fit-for-modelling-purposes, disease-specific data is currently lacking from traditional sources like hospitals. This data gap can be addressed by investing in wearable and implantable sensor technologies capable of providing real-time patient data. Micro-physiological system technologies present another avenue for generating high-quality human-relevant data by combining microfluidics with embedded sensors. These technologies provide a platform for investigating adverse drug effects and modelling specific organ functions, offering valuable insights for the development of VHTs.
3. **Evidence Generation for VHT Development:** The development of robust and trustworthy VHT technologies hinges on the generation of comprehensive clinical, experimental, and digital evidence to support model building and validation. This process requires well-designed studies and clinical trials that first assess the technical validity of VHT solutions, followed by their clinical domains such as safety, effectiveness, and usability. A deep understanding of the unmet needs of VHT early adopters and how these technologies will be implemented to address those needs is essential for generating relevant and impactful evidence. VHT technologies can be used to develop synthetic and open-access validation datasets, ensuring transparency, reproducibility, and accessibility throughout this process. This evidence generation process will be instrumental in building confidence among stakeholders, paving the way for wider adoption of VHT in healthcare.
4. **Credibility is Key to Building Trust in VHT Technologies:** To foster trust and confidence in VHT technologies, establishing their credibility is essential. This involves a multi-faceted approach encompassing rigorous and transparent safety assessments and peer review/validation of methods to ensure scientific rigor and transparency. Clear procedures for verification and validation, for a well-defined question of interest and context of use, must be developed, drawing upon established ontologies to provide a standardized framework for credibility assessment. Uncertainty quantification plays a vital role in understanding the limitations of VHT predictions and enabling risk-informed decision-making. Addressing potential refutations and providing clear evidence to support the credibility of VHT findings will be crucial for gaining acceptance from stakeholders.

PART 3:

REALISING THE VHT – INFRASTRUCTURE

15 Infrastructure

15.1 A trinity of software

As outlined in PART 1 of this roadmap, the realisation of the VHT requires collaborative efforts at a European level, including the establishment of a dedicated VHT infrastructure. The infrastructure will play a key role in pooling resources and expertise, data sharing and interoperability, credibility and trust, as well as catalyse economic sustainability. The term Infrastructure refers to the **trinity of software Catalogue, Repository and Platform**, that together create a one-stop shop to discover, share, design and use DTs in health and care. In general, the VHT Platform will accomplish data analyses, simulation and visualization by processing, managing and interacting with Research Objects - data, tools, services, workflows - found under the VHT Catalogue and (possibly but not exclusively) stored in the VHT Repository. The realization of such implementation comes by delivering the VHT Platform in a ‘Platform as a Service’ (PaaS) way, accomplished by technological solutions focused on container and cloud-based deployment and configurations.

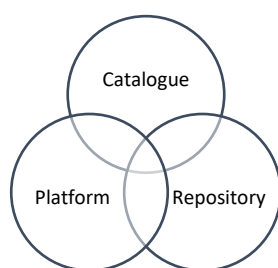


Figure 15: schematic overview of the main components of the VHT infrastructure.

The **infrastructural requirements** of the VHT platform are a **direct consequence of the requirements** of the underlying computational models and data storage facilities. In order to provide the services envisioned in this roadmap (*cfr.* PART 2), both the repository and platform need to be able to access a variety of computational resources, and operate across organisational boundaries, in an international and intersectoral context. Enabling such operation calls for resolution of a range of issues of legal, organisational and technical nature. A useful primer to European infrastructural standardisation efforts is provided by the Rolling Plan for ICT Standardisation. This is a live document, curated by the European Commission in collaboration with the European Multi-Stakeholder Platform (MSP) on ICT Standardisation. It lists all the topics identified as EU policy priorities where standardisation, standards, or ICT technical specifications may play a key role in the implementation of the policy.

15.2 Strategic pillars

What we describe here is the general vision of what, in the long term, the VHT infrastructure should be. The **VHT infrastructure** should be developed around four strategic pillars considered essential for the global effort to drive forward **VHT** development.

1. Distributed/federated architecture
2. Governance
3. Openness
4. User engagement

15.2.1 Pillar 1: Distributed/federated architecture

The heart of the VHT infrastructure will consist of a distributed/federated platform. The different services of the VHT platform will be semantically mapped to different types that occur within the context of the VHT. The platform will include centralised **core elements**, *i.e.*, the back-end elements required to run the platform. The second level will include the **platform and science-specific services**, *i.e.*, generic platform elements necessary to the end-users (*e.g.*, wiki, collaborative documents), scientific services like the repository, tools for running the workflows, *etc.* These might include services for semantic re-annotation or services to promote resources along the credibility axis. The last level,

domain-specific services, is also end-user facing and includes all the federated services relevant to the VHT.

A future benefit of administratively treating **domain-specific** products and **services** separately from **Platform and Science specific services** is that it facilitates an easy onboarding path for more domain-specific tools and services and even tools from other scientific domains. The onboarding is further facilitated by the federated – more flexible – nature of the VHT platform. Component owners will enjoy full *autonomy* regarding the services and applications they provide. New services will be developed, mature, and be provided as federated services that will be part of the VHT platform, following integration/quality/interoperability requirements.

The VHT infrastructure needs to be “**as open as possible, as close as necessary**”, *i.e.*, very accommodating of various existing components, formats, and protocols, but at the same time, it needs to provide a unified and intuitive user experience that does not expose all the “sharp edges” of the underlying machinery. This follows the Robustness Principle (Postel’s Law) made famous during the specification of the TCP (transmission control protocol): “*be conservative in what you do, be liberal in what you accept from others*”¹⁶⁸.

The implementation as a **distributed/federated** platform will be more **flexible** and **adaptable** to changing requirements and user needs. Different entities can choose which technologies and standards to use and evolve their systems independently. Moreover, its distributed nature will allow the platform to easily **scale** to handle increased traffic or user demands by adding more servers or nodes and to be deployed across **multiple locations**, allowing for a wider reach and better performance for users in different regions. Given the distributed nature of the storage services archiving data and the computing services elaborating on them, they need to be connected by high-speed geographic networks.

A federated/distributed infrastructure can also promote **interoperability** between different systems, allowing users to communicate and share data across different platforms, and fostering innovation if proper standards and interfaces are defined and enforced. At the same time, it also improves the **performance** of the platform by reducing latency and increasing bandwidth. However, it may also require more coordination and **governance** to ensure interoperability and maintain quality.

15.2.2 Pillar 2: Governance

Governance is essential since it defines the management structure, roles, and decision-making procedures. The **governance framework** will identify/establish the policies, governing roles and responsibilities (admin, provider, and user profiles) and decide the standards to adopt. It will also ensure the clarity of the business models and access policies (tiers, pricing policies, commercial agreements).

15.2.3 Pillar 3: Openness

Openness in the VHT infrastructure will allow users and developers to share their work and collaborate with others, bringing together people with different backgrounds and expertise, leading to more diverse perspectives and insights, reducing, at the same time, duplication of effort and resources by allowing users and developers to build on each other’s work. Note that ‘open’ is not equal to ‘free’. The research findings will be available to the entire ecosystem: the scientific community, industry, policymakers, and other stakeholders. An **open VHT infrastructure** will also facilitate the sharing of data and resources, encouraging, at the same time, uniformity of protocols and formats, standardised wherever possible and standardisable in hopefully all other cases (*cfr.* PART 4). As an important element to guarantee openness, software shall be required to be licenced in a way that permits open access to the source code for those software entering the VHT infrastructure. This does not prevent transfer of rights for commercialisation from contributors to companies if necessary within business models. Additionally, agreements can be negotiated with commercial companies to provide a number of licenses to use some of the shared pipelines within the VHT.

15.2.4 Pillar 4: User engagement

A **user-friendly and visually appealing platform** will facilitate the ecosystem’s interaction with the VHT infrastructure. The design needs to be user-centred, and user surveys will be used to gather insights that inform the design and functionality in a co-creation process. The layout needs to be easy to navigate.

¹⁶⁸ Postel, Jon, ed. (January 1980). Transmission Control Protocol. IETF. <http://doi.org/10.17487/RFC0761>.

The platform needs to have a **dashboard** that will guide the user through the different offerings of the VHT, allowing the users to customise their profiles and add the services and tools that are useful to them. Clear documentation needs to be available, not only for the main functions of the platform but also for the different services. The platform will combine tools and services from different sources, all valuable for developing the VHT. Users will use the platform also for collaboration and interaction with other users. The ecosystem will provide user incentives for sharing data, models, or other content (*cfr.* PART 5), while it will consider the feedback from the users, working to improve itself. Finally, user engagement should be measured and analytics tools will be used to monitor user behaviour and identify improvement areas.

Box 24: Success story – the 12 LABOURS project.

Success story: Developing a Digital Twin Platform linked to clinical workflows

Status: used in clinical research

Website: <https://www.auckland.ac.nz/en/abi/our-research/research-groups-themes/12-Labours.html>

The clinical translation of personalized computational physiology workflows and Digital Twins has the potential to revolutionize healthcare by enhancing our understanding of individual physiological processes and identifying changes that may lead to serious health outcomes. However, the absence of a unified infrastructure for developing these workflows and Digital Twins has hindered the realization of this transformative vision. The **Auckland Bioengineering Institute's 12 LABOURS project** seeks to address these challenges through the development of a **Digital Twin Platform**¹⁶⁹. This platform enables researchers to develop and personalize computational physiology models based on individual health data within clinical workflows, thereby facilitating more efficient personalised clinical treatments and home-based care. The 12 LABOURS project, builds upon over two decades of research from the Physiome Project, in synergy with international projects such as SPARC (US-NIH) and VITAL (EC-Horizon Europe).

The platform's capabilities are developed through a range of demonstrators, covering major organ systems. These demonstrators show the platform's capabilities using publicly available data and an existing automated computational physiology workflow designed to assist clinicians in diagnosing and treating diseases. The platform **enables the discovery and exploration of data while presenting workflow results in clinical reports via a web portal**. Current enhancements of the platform focus on integrating it with health systems and remote monitoring devices, such as wearables and implantables, to support home-based healthcare. Additionally, by combining outputs from multiple workflows applied to the same individual's health data, the platform will enable the creation of personalized Digital Twins.

The 12 LABOURS Digital Twin Platform provides several significant **clinical and scientific benefits**. First, it allows researchers to more efficiently conduct clinical trials to evaluate the efficacy of computational physiology workflows, facilitating their clinical translation. Second, it enables the reuse of primary and derived data from these workflows to generate novel workflows. Finally, it supports the creation of personalized Digital Twins by integrating outputs from diverse computational physiology workflows, thereby advancing personalized medicine.

¹⁶⁹ <http://doi.org/10.1109/EMBC40787.2023.10341138>

16 Catalogue and repository

16.1 Catalogue & repository: conceptual representation

The VHT repository is envisioned as a space where VHT assets can be uploaded, stored, discovered and accessed by users. Assets refer not only to data objects (*e.g.*, input files or results of VHT simulations), but also to the models objects and other resources as previously defined. The European community has had some success in building such cross-organisational repositories – as evidenced by the EOSC Marketplace¹⁷⁰, the EUDAT project or the Fenix infrastructure¹⁷¹ – and the VHT repository framework should build upon the achievements of such initiatives, while also extending their capabilities to provide support for a VHT-specific metadata model.

In this sense, the VHT repository/catalogue should provide support for the elements listed below, schematically depicted in figure 6.

- **Catalogue services**, enabling discoverability of resources;
- **Secure role-based access** (see next chapter for a discussion of user roles);
- **Dedicated storage resources** for unstructured binary data, structured databases and model repositories (with CASE-specific features, such as those provided by modern code versioning tools).

Another crucial aspect for any actionable VHT repository is the provision of a consistent set of programmer's interfaces (APIs) through which content could be accessed programmatically rather than directly by human actors. Such a feature would facilitate development of automated computational workflows and enable large-scale VHT simulations.

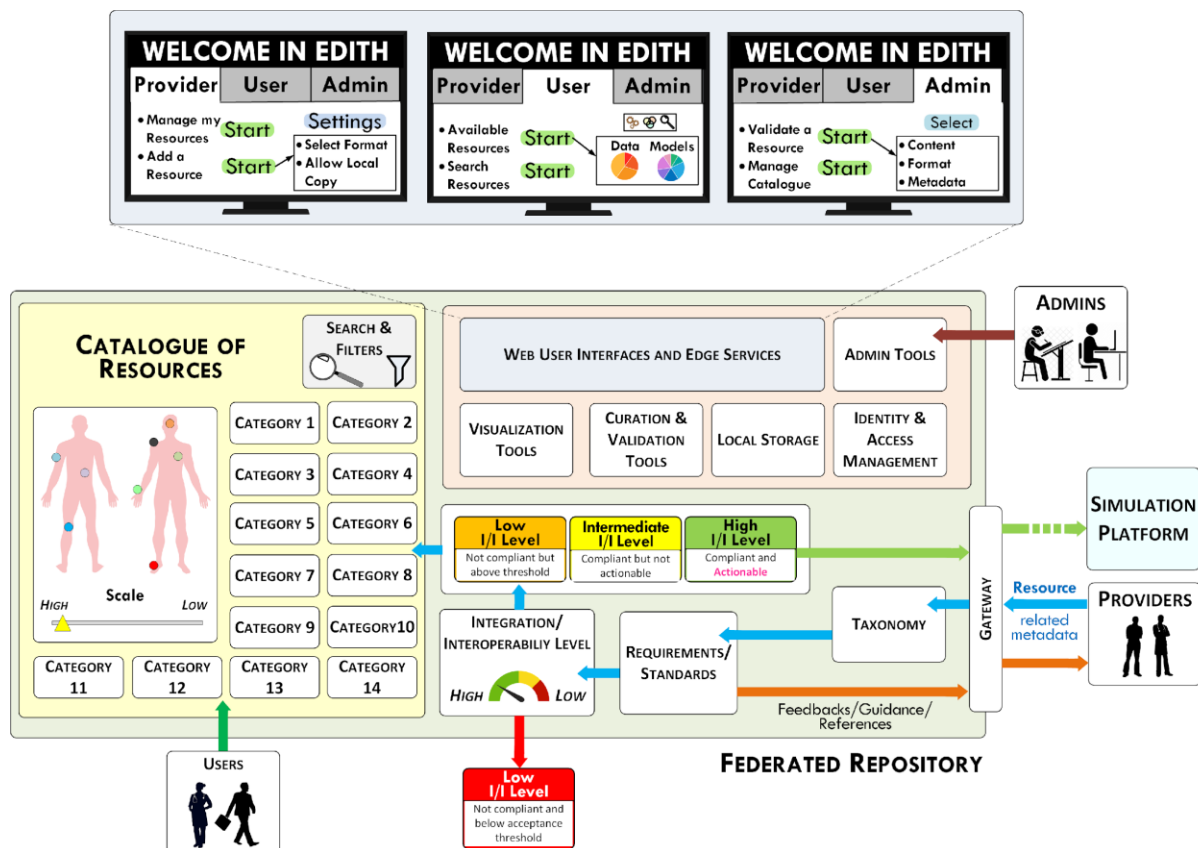


Figure 16: schematic representation of the repository and the catalogue depicting its essential components.

To better appreciate possible potentials and pitfalls in the development of the VHT that needed to be addressed in this roadmap, different user perspectives were investigated. The example below presents the possible journey from the provider point of view (*i.e.*, who wants to share a resource within the

¹⁷⁰ <https://marketplace.eosc-portal.eu>

¹⁷¹ <https://fenix-ri.eu/infrastructure/resources/available-resources>

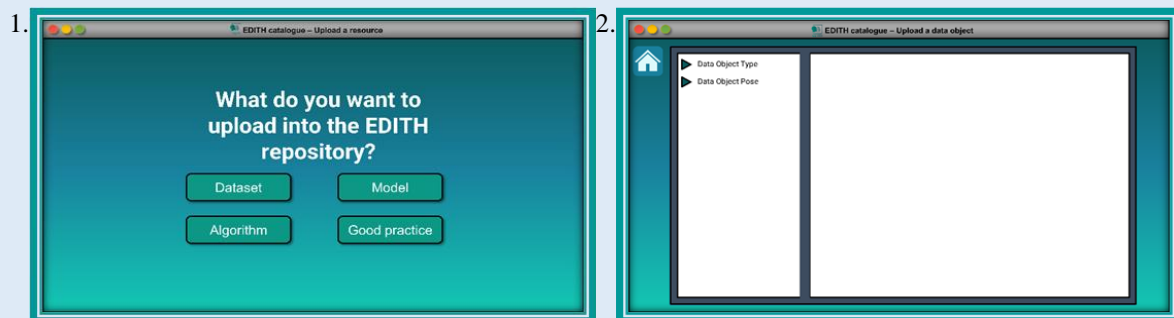
EDITH platform) and the user point of view (*i.e.*, who wants to use one or multiple resources available in the EDITH platform).

Box 25: EDITH development – mock-up for uploading a resource in the repository

EDITH development : mock-up user interaction with infrastructure – uploading a resource

Upon accessing the infrastructure, an authentication is requested. Once authentication has taken place, the user is asked whether the action to be carried out concerns uploading or using a resource. The mock-up illustration below concerns uploading a resource.

1. Once the “upload a resource” option is selected, the provider is asked which type of resource will be uploaded. The provider can choose among dataset, model, algorithm, and good practice.
2. The provider is asked to insert information about the resource to be uploaded. The dataset must be annotated with sufficient metadata, which includes information on the Data Object Type (DoT) and Data Object Pose (DoP). The DoT comprises information about semantics, syntax and accessibility. The DoP comprises information about the dimensions of space, time, credibility and clustering.



3. The DoT definition starts from the Semantics, representing the basic information of the dataset. In this case, an atlas and a scale is presented, asking the provider to select respectively the dimension and the physiological system the dataset is referred to.
4. Afterwards, the provider is asked if the resource is based on an ontology. If yes, the provider can select the ontology used from a pre-defined list or, otherwise, input the name of the ontology used if not present in the list. If the dataset's variables are already set following an ontology present in the VHT platform, the provider has to simply select the ontology and move forward.
5. The second element of the Data Object Type corresponds to the syntax definition. Also in this case, the provider is asked if the resource is standardized and, if yes, which standard has been implemented. As seen for the Semantic annotation, if the resource follows a standard, a set of standards will appear and the provider can select the standard used. Here there are two different scenarios that may happen:
 - a. The standard used for the resource is the one adopted for the whole VHT platform.
 - b. The standard used is present in the list, but it is not the one adopted for the whole VHT platform. In this case, an automatic mapping procedure will take place for the conversion from the current standard to the VHT platform one.
 - c. The standard used for the resource does not follow a standard among the ones listed. In this case, a guided procedure for standard mapping will have to take place. The latter will also happen in case the dataset does not follow any standards.
6. The last point for the Data Object Type definition involves the definition of possible license in using and sharing the resource being uploaded. The provider has also the possibility to define his/her own terms of conditions for that specific resource.



Once the DoT information has been inserted, the provider is asked to provide information regarding the DoP.

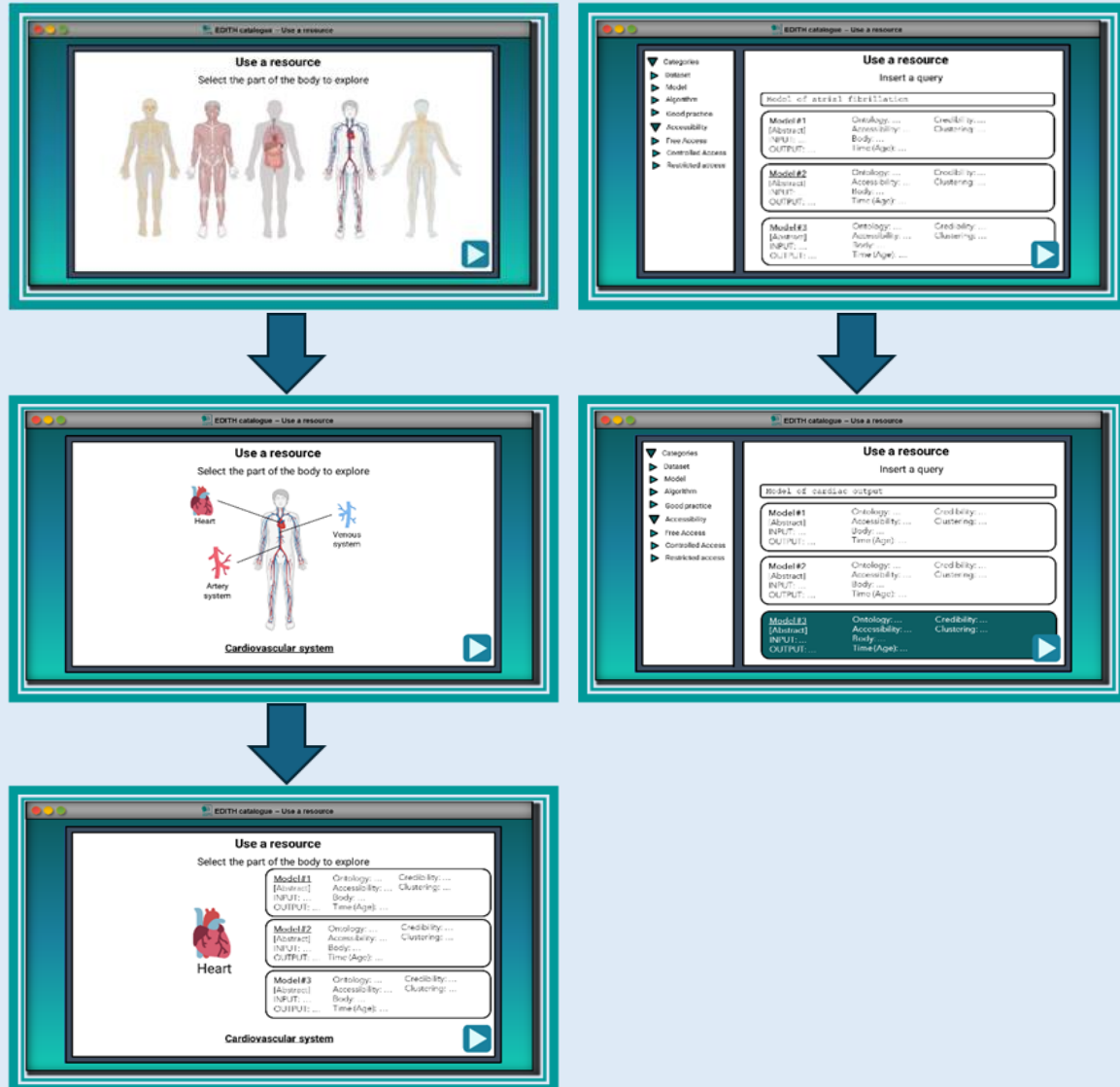
7. The Body interface allows defining the anatomical characteristics of the resource being uploaded. For those resources having a specific location in the human body, appropriate transformations must be defined as explained in PART 2. Thus, through the Body interface the provider may select the most appropriate rigid transformation, both for a single and for a multi-body object, and the elastic transformation.
8. In this mock-up interface, Time is defined as the age of the involved participants to the resource. Here a normalized time axis is presented, spanning from 0 (birth) to one (death).
9. The Credibility interface provides the possibility of defining the level of trustworthiness of the resource being uploaded. It spans in a range from 0 (no credible) to 1 (the credibility is certified by a third party).
10. The Cluster property allows to define that the resource comprehends different space and/or time scales within the human body. Also in this case, the cluster property can be defined on a range spanning from 0 (no clustering) to 1 (homo sapiens).



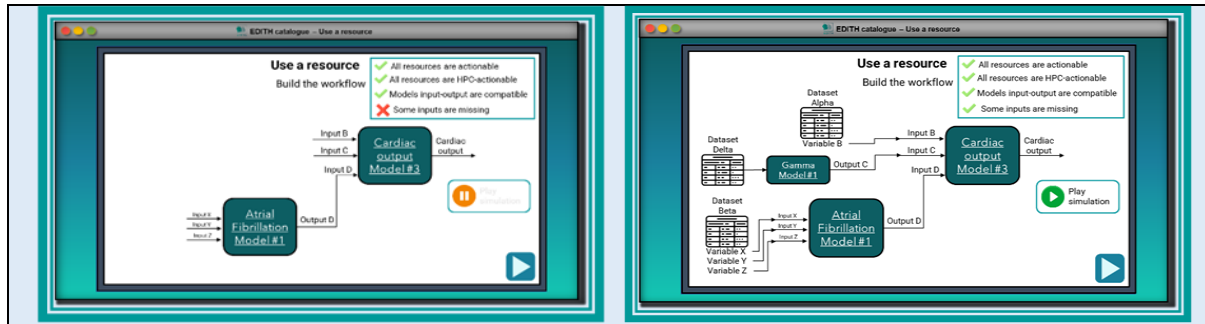
Box 26: EDITH development – mock-up for preparing a model in the repository

EDITH development : mock-up user interaction with infrastructure – preparing a model

Let's hypothesize that the user wants to model the cardiac output in presence of atrial fibrillation. The user can select the model, or multiple models to obtain the final outcome, through a graphical user interface, by selecting the anatomical district the models are related to (left), or by typing it in a search query (right).



Once the models are selected, the user can build the workflow in an operational environment within the EDITH platform by connecting inputs and outputs of the selected models. The environment conducts a technical validation to inform the user whether the selected models are ready to be run or not. Options about the actionability of the models and the compatibility of the selected models will be automatically carried out. The environment will inform also about the possibility of some models lacking of inputs (left, below). In that case, the user might use the already uploaded datasets for those inputs – if compatible (right, below). Once the technical validation provides a positive result, the simulation can be run.



16.2 State of the Art

In the evolving landscape of medical care and personalised medicine, understanding and having an overview of the **state-of-the-art** technologies and initiatives regarding health care is crucial for staying at the forefront of scientific and technological progress. In order to depict and structure the details of the VHT Catalogue, Repository and Platform, an exploration of various **tools**, **platforms**, **environments** as well as **data repositories** and **data type references** was essential.

16.2.1 Initiatives inside the EU

The implementation of **access to personalised data** within the VHT can benefit from ongoing research in terms of personal health data access, such as in the development of personal Health Data Pods¹⁷² (POD: personal online data) in EU member states. These data pods could contain a wide array of health data types ranging from hospital records, such as MRI scans, ECG results, personal genome data and diagnoses, to pharmacy data, and even data captured from wearable devices. The VHT infrastructure will need to foster an environment where personal health data can be securely accessed and effectively utilised (primary use) while retaining patient consent.

Regarding **personal genome data**, Denmark has made remarkable development in personalised medicine partnership¹⁷³ via a Federated Trusted Research Environment. Additionally, the launch of a desktop workbench, QIAGEN¹⁷⁴, for whole genome sequencing represents a notable achievement. This initiative incorporates important elements, including tiered access models based on users accepting various licenses. These licenses empower users to select from an array of tools and workflows that align with their scientific needs. Moreover, at the EU level, the “Beyond 1 Million Genomes” (BIMG) initiative¹⁷⁵ aims to deliver personalised medicine on a transnational scale. Although the complexity of this endeavour presents challenges that must be addressed, the potential to access and link phenotypic data has the power to leverage the healthcare systems. Some of the challenges refer to various levels of maturity in genomic medicine programs among different countries, diverse healthcare system infrastructures and distinct processes. Towards the realization of BIMG’s goals, a collaborative effort is established in partnership with 1+ Million Genomes (1+MG) Initiative’s working groups.

In the context of **secondary use of health data**, the Towards European Health Data Space (TEHDS) initiative successfully concluded the groundwork¹⁷⁶ for the European Health Data Space (EHDS). This action focused on exploring various technical options designed to facilitate the secondary use of health data within the EU. These technical options emphasize a secure processing environment (SPE), where data processing takes place after a data permit has been granted. Throughout this initiative, a centralised “one stop shop” approach was introduced, enabling users of health data to submit their permit requests to a single designated location responsible for its management, even if data is distributed among multiple data holders. Following the conclusion of this phase, the HealthData@EU¹⁷⁷ initiative was

¹⁷² <https://we-are-health.be/en>

¹⁷³ Spinner J., Danish National Genome Center, Lifebit form personalized medicine partnership, 2022. <https://www.outsourcing-pharma.com/article/2022/06/01/danish-agency-lifebit-form-personalized-medicine-partnership>

¹⁷⁴ <https://corporate.qiagen.com/English/newsroom/press-releases/press-release-details/2023/Danish-National-Genome-Center-selects-QIAGEN-for-variant-interpretation-in-oncology-genome-sequencing/default.aspx>

¹⁷⁵ <https://bimg-project.eu/work-packages/wp5>

¹⁷⁶ <https://tehdas.eu/results/tehdas-proposals-for-the-implementation-of-ehds-technical-infrastructure/>

¹⁷⁷ <https://ehds2pilot.eu/>

established to serve research, innovation, policy making and regulatory purposes. This initiative will create a network infrastructure connecting data platforms across EU Member States. It will also provide guidelines for data standards, quality, security and cross – border data transfer, ensuring a seamless and standardised approach to the secondary use of health data within the EU.

16.2.2 Initiatives outside of the EU – Trusted Research Environments

Outside of the EU, within the UK context, a comprehensive and well-structured research framework for personal data has already been established. For a long period of time, Data Safe Havens also known as Trusted Research Environments¹⁷⁸, have played a transformative role in ensuring secure access, analysis and retrieval of data outputs. Delving into specific UK Trusted Research Environments (TREs), examples such as OpenSafely¹⁷⁹, NHS England¹⁸⁰, Genomic England TRE¹⁸¹ and SAIL¹⁸² stand out. All TREs adhere to “**five (5) Safe principles**”¹⁸³ including Safe People, Projects, Settings, Data and Outputs.

- **Safe People:** certified users are required evidence of research careers, affiliations with recognised institutions, and a commitment to complete data-safe research training.
- **Safe Projects:** a project proposal must demonstrate the public benefits it will deliver.
- **Safe Settings:** these settings provide technical solutions in terms of secure analyses, storage and access. While some TREs necessitate physical attendance in strictly monitored workstations (SeRP¹⁸⁴), others use Virtual Desktop Infrastructures (VDIs) for a more flexible approach. VDIs come with a pre-defined suite of software, including data analysis tools like R, Python, SPSS, STATA, as well as including whole Integrated Development Environments (IDEs) from eclipse and Databricks, to Jupyter notebooks. As seen in OpenSafely and Genomic England TREs, users can also access interfaces to create and execute Machine Learning algorithms, as well as workflows written in Nextflow¹⁸⁵ and Workflow Descriptive Language (WDL)^{186,187} and executed in the Cromwell workflow engine.
- **Safe Data:** the concept is reducing re-identification risk through pseudonymization, encryption and decryption of linkage keys.
- **Safe Outputs:** human review and manual intervention are required. In almost every TRE, Airlock¹⁸⁸ is used in order to whitelist types of outputs that can be exported.

In addition, a tremendous initiative is underway to incorporate Safe Return and Safe Computing principles, each posing unique challenges that require careful consideration.

- **Safe Return** places a particular challenge related to ethical approvals, patient consents and the establishment of a certified data path.
- **Safe Computing** focuses on addressing technical complexities. It is worth mentioning that Amazon Web Services (AWS) is widely used across the majority of TREs¹⁸⁹, both for secure data storage both at rest and in transit, as well as for analysis of personal data.

There is a growing movement enhancing principles towards **federated Trusted Research Environments**¹⁹⁰, allowing for collaboration and secure data access in a federated manner. This approach is leading the way to federated data sharing and research opportunities across diverse TRE platforms. With this initiative, the five (5) Safe principles are reinforced with open standards for Authentication and Authorization (such as GA4GH Passports and Visas¹⁹¹), ways of federated analyses

¹⁷⁸ https://www.hdruc.ac.uk/wp-content/uploads/2021/09/HDRUK_TRE-One-Pager.pdf

¹⁷⁹ <https://www.opensafely.org/>

¹⁸⁰ <https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england>

¹⁸¹ <https://www.genomicsengland.co.uk/research/research-environment>

¹⁸² <https://saildatabank.com/>

¹⁸³ <https://doi.org/10.5281/zenodo.5767586>

¹⁸⁴ <https://serp.ac.uk/about-serp/>

¹⁸⁵ https://re-docs.genomicsengland.co.uk/small_variant/

¹⁸⁶ https://re-docs.genomicsengland.co.uk/sv_cnv/

¹⁸⁷ https://re-docs.genomicsengland.co.uk/sv_cnv1/

¹⁸⁸ <https://www.airlockdigital.com/>

¹⁸⁹ <https://docs.aws.amazon.com/whitepapers/latest/nhs-cloud-security-guidance-using-aws/nhs-cloud-security-guidance-using-aws.html>

¹⁹⁰ <https://dareuk.org.uk/wp-content/uploads/2023/04/DARE-UK-Federated-Architecture-1-Initial.pdf>

¹⁹¹ https://www.ga4gh.org/news_item/ga4gh-passports-and-the-authorization-and-authentication-infrastructure/

through workflow and tool engines like GA4GH WES¹⁹²/ TES¹⁹³ standards, and the adoption of common data standards, such as the Observational Medical Outcomes Partnership (OMOP)¹⁹⁴.

16.2.3 Repositories and data management tools

Various **repositories** hold objects spanning from datasets to tools and workflows. Such examples are Zenodo¹⁹⁵ and the openEHR Clinical Knowledge Manager (CKM)¹⁹⁶. The former is a general-purpose open repository developed under the European OpenAIRE¹⁹⁷ program and operated by CERN. It allows researchers to deposit research papers, data sets, research software, reports, and any other research related digital artefacts. The latter is an international, online clinical knowledge resource. It is a powerful collaboration tool to develop, manage, visualise, review and publish clinical data models. In addition, HealthData.gov¹⁹⁸ brings together high-value datasets, tools, and applications, providing health-related data to assist researchers in problem-solving. It houses thousands of health-related data sets from USA Department of Health and Human Services agencies with new datasets being regularly added.

Box 27: Success story – building the Human Cell Atlas.

Success story: large scale data sharing and community building

Status: established legal entities in UK and the Netherlands

Website: <https://www.humancellatlas.org/>

The **Human Cell Atlas (HCA)** is a global scientific initiative aimed at creating detailed, comprehensive reference maps of all human cells—the fundamental building blocks of life. By cataloguing and understanding the unique types and states of cells in the human body, the HCA seeks to revolutionize biology and medicine. These maps are designed to serve as **foundational resources for exploring the cellular basis of health**, as well as the molecular and cellular changes associated with diseases such as cancer, autoimmune disorders, and infectious diseases. Launched in 2016, the HCA has grown into a vast consortium of over 3,200 members representing more than 1,700 institutions across 99 countries, fostering a highly collaborative and interdisciplinary research environment. Researchers in the consortium employ cutting-edge technologies, including single-cell transcriptomics, proteomics, and spatial genomics, to chart the molecular profiles, spatial organization, and interactions of cells within tissues and organs. This work is advancing our understanding of cellular diversity and function, enabling transformative insights into human biology and disease mechanisms. Central to the HCA's mission is the **HCA Data Portal**¹⁹⁹, a state-of-the-art open-access platform that serves as the repository for the consortium's extensive datasets. The portal plays a crucial role in organizing, sharing, and enabling the analysis of data collected from thousands of research projects around the world. As of now, the portal hosts data on approximately 63 million cells, collected from 9,200 donors and contributed by 835 laboratories across 478 projects. These datasets include high-resolution information about individual cells' genetic, transcriptomic, and epigenetic profiles, as well as their spatial organization within tissues. Researchers can use the portal to explore multi-omic data spanning diverse biological systems, including organs such as the heart, lung, and liver, as well as networks like the immune and nervous systems. By providing researchers with a unified platform for data access and exploration, the HCA Data Portal accelerates discovery and fosters collaboration, helping scientists worldwide tackle complex questions about cellular function, human development, and disease progression. This resource is designed to evolve continuously, integrating new datasets and tools as they become available, ensuring it remains a vital resource for the global biomedical research community.

The HCA's endeavours are supported by a **diverse array of funding sources globally**. Notable contributors include the Chan Zuckerberg Initiative (CZI), which has provided substantial funding to support collaborative science teams and seed networks for the HCA. Additionally, the Wellcome Trust has allocated £7 million to UK scientists to scale up the HCA initiative. Other significant funders include the Helmsley Charitable Trust, which has partnered with CZI on funding mechanisms to support the HCA community, and the European Union's Horizon 2020 program, which has funded pilot actions to build the foundations of the HCA. Through

¹⁹² <https://www.ga4gh.org/product/workflow-execution-service-wes/>

¹⁹³ https://www.ga4gh.org/news_item/ga4gh-tes-api-bringing-compatibility-to-task-execution-across-hpc-systems-the-cloud-and-beyond/

¹⁹⁴ <https://www.hdruk.ac.uk/wp-content/uploads/2021/06/210622-Recommendations-for-Data-Standards-2021-Interim-Paper.pdf>

¹⁹⁵ <https://zenodo.org/>

¹⁹⁶ <https://ckm.openehr.org/ckm/>

¹⁹⁷ <https://www.openaire.eu/>

¹⁹⁸ <https://healthdata.gov/>

¹⁹⁹ <https://data.humancellatlas.org/>

these collaborative efforts and diverse funding sources, the Human Cell Atlas aims to revolutionize biomedical research and clinical practice by providing an unparalleled understanding of human cellular biology.

The Comprehensive Knowledge Archive Network (CKAN)²⁰⁰ stands out as a **data management system** that offers a platform for cataloguing, storing and accessing datasets. Some key features include an API for accessing the core functionality, a Datastore extension for structured data storage, and the ability to customize the platform to meet specific data portal needs. CKAN also offers a robust search experience, enabling quick keyword search and filtering. Alternatively, InvenioRDM²⁰¹ represents an open-source, community-driven research data management (RDM) repository. It has a modern web architecture and adheres to standards that facilitate easy deployment, maintenance, and usage. Its development includes a wide range of features aimed at streamlining good data practice and enhancing the value throughout the research lifecycle. OpenClinica²⁰² has been known as the world's first commercial open source clinical trial software, serving for the purpose of clinical data management (CDM) and electronic data capture (EDC). OpenClinica is deeply embedded in clinical trials and various types of clinical research domains. This modern electronic data capture software typically operates via a thin client on a web-based platform, enabling the electronic capture of clinical trial data or the option to capture data in paper form, which can later be transcribed into the EDC system. Thin clients primarily rely on a web server and can access and utilize the EDC software through internet connectivity. Data collection sources for clinical trial patients or subjects are typically hospitals or clinics. In the context of an EDC system for a study, nurses or other designated study coordinators at the site are usually responsible for data entry. The individuals responsible for reviewing and electronically signing the data are the site investigators and physicians overseeing the patient's care and data²⁰³.

Moving toward the neuroscience research area, the EBRAINS Knowledge Graph²⁰⁴ is a platform that consolidates information from various areas of brain research, linking research data with software designed for data analysis. At its heart, the EBRAINS Knowledge Graph (KG) employs a **graph database** to connect neuroscientific research, data, tools, services across different modalities, utilizing the openMINDS metadata model (metadata framework for neuroscience graph databases containing metadata models), libraries of controlled terminologies, brain atlases and common coordinate spaces. Over the last few months, EBRAINS KG also played an important role in enhancing shareability, findability and completeness of Standardised Workflows²⁰⁵. Key element of EBRAINS KG is the adoption of prospective metadata of two individual Standardised Workflows^{206,207}.

Box 28: Success story – Digital Twins of the Brain

Success story : Digital Twins of the brain

Status: *evaluation in clinical context*

Website: <https://www.thevirtualbrain.org/tvb/zwei>

The human brain is the most complex information processing system known today. It is organized across a wide range of scales; with strongly interwoven architectural principles spanning from the level of molecules and cells up to the whole organ which generate the spatially and temporally distributed activity patterns that determine our perceptions and behaviour.

Digital Twins of the human brain are of utmost importance for basic research in clinical applications. For example, while deep brain stimulation is an essential explorative tool to localize epigenetic zones for the diagnosis of epilepsy, the number of regions that can be explored is strongly limited by the invasive nature of implanting the necessary electrodes. At the same time, findings cannot be easily translated from patient to patient since epilepsy is a highly patient-specific disease.

²⁰⁰ <https://ckan.org/>

²⁰¹ <https://inveniosoftware.org/products/rdm/>

²⁰² www.openclinica.com

²⁰³ <https://sollers.edu/medical-programs/clinical-trial-management-course/>

²⁰⁴ <https://www.ebrains.eu/data/find-data/find-data>

²⁰⁵ <https://www.humanbrainproject.eu/en/science-development/scientific-achievements/deliverables/third-specific-grant-agreement/>

²⁰⁶ HBP Showcase 3: Workflow. <https://search.kg.ebrains.eu/instances/bd71c7c0-c2bb-454a-8d18-12ef96f45cdd>

²⁰⁷ HBP Showcase 4: Workflow. <https://search.kg.ebrains.eu/instances/86efc59a-b4d1-443d-a470-dd27ddc4465b>

Here, a **personalized simulation of brain activity** with realistic anatomical constraints can have a major impact. Such personalized brain models can be realized using the TVB (the Virtual Brain²⁰⁸) simulation framework, which allows to model patient specific brain activity in 3D based on various anatomical priors that include anatomical brain regions with their mutual connection strengths and characteristic regional attributes.

However, *in vivo* measurements obtained from the individual patient cannot provide a rich characterization of the **detailed multiscale brain organization**. Therefore, the simulation needs to connect with a comprehensive **brain atlas** which provides detailed measurements of functionally distinct brain areas, their connectivity, their variability across different individuals, and their characteristic molecular and cellular properties. Such a combination of personalized brain simulations with detailed multilevel atlas information has been realized by coupling TVB with the multilevel human brain atlas in EBRAINS²⁰⁹ through the *siibra* software²¹⁰. The atlas framework uses the Julich-Brain probabilistic cytoarchitectonic maps²¹¹ as a reference, and integrates the BigBrain model as a 3D anatomical model with a spatial resolution of 20 micrometres isotropic that allows it to capture the cellular architecture with cortical layers and subcortical nuclei. This way, detailed anatomical constraints can be leveraged for building personalized whole-brain model soft brain activity in TVB to simulate brain-wide activity.

The system is currently being **evaluated in clinical environments** at Aix-Marseille university and the neurosurgery unit of the university hospital Düsseldorf.

Finally, in the UK context, the **Health Data Research (HDR) Gateway** serves as the entry point for researchers and innovators seeking access to UK health datasets. The HDR Gateway itself does not store, hold, or process any patient or health data. Instead, its purpose is to assist users in locating datasets originating from and stored within specific TREs and streamline the process of requesting access. A notable feature of the HDR Gateway is the Cohort Discovery function. This function enhances data discovery by enabling users to query data references across multiple TREs, providing information on the number and type of datasets within each TRE. These queries are conducted against pseudonymized (de-identified) datasets securely hosted behind the firewalls of data custodians. This feature holds significance as it empowers users to decide, based on the retrieved information, whether they wish to request access to specific datasets and TREs and assess the potential public value of their findings.

Box 29: Success story – Data Collection Initiative in Germany

Success story : the German Medical Informatics Initiative

Status: *being tested in clinical use cases*

Website: <https://www.medizinformatik-initiative.de/en/start>

The Medical Informatics Initiative (MII) in Germany, funded by the Federal Ministry of Education and Research (BMBF), aims to establish a robust data infrastructure throughout Germany. A central component of this initiative is the creation of Data Integration Centers (DICs) at university hospitals, which consolidate data available at the hospitals, ensuring quality and compliance with data protection regulations. These DICs serve as decentralized technical and organizational hubs, enabling the collection, integration, and processing of diverse healthcare data.

This approach facilitates the effective use of data in both healthcare and research, promoting interoperability and reusability. A harmonized data format ensures that researchers can access and analyse data from multiple university hospitals through a single query within a unified legal framework.

In conclusion, with the information from this section in mind, in order to realize the ambition of the VHT, it is imperative that the VHT embraces **standardization methods** not only in data analysis but also in data description. Following the example set by EBRAINS KG with its Standardized Workflows approach, the VHT can enhance the efficiency and reproducibility of the ecosystem's research efforts. Furthermore, an integral aspect of the VHT's success lies in the enhancement of an **authentication and authorization infrastructure** (AAI) by accounting and tiering in the scope of user profiles and user roles. The former offer a secure way of authenticating users and make strong connections with affiliations and research areas, whereas the latter offers a specific level of access to part of the systems. By providing tailored access to authenticated individuals and securing their engagement within the

²⁰⁸ <https://www.thevirtualbrain.org/tvb/zwei>

²⁰⁹ <https://ebrains.eu>

²¹⁰ <https://siibra-python.readthedocs.io>

²¹¹ <http://doi.org/10.1126/science.abb4588>

appropriate datasets and secure environments, a robust framework can be established that promotes both accountability and data security. More comprehensive information about the VHT Repository, Catalogue and Platform can be found in the following sections.

16.3 Catalogue and Repository within the VHT Infrastructure

While this chapter primarily focuses on the VHT Repository, it is essential to provide a concise overview of the entire VHT Infrastructure's design as well as the repository's position within it. The VHT Infrastructure comprises a trinity of software: Catalogue, Repository and Platform. The table below outlines the primary functions of each part, as they were defined during the EDITH-CSA project, taking into account user requirements, collaborative meetings, and discussions within and beyond the EDITH consortium. More specific requirements for the VHT Catalogue and Repository are provided in the next section.

Table 1: EDITH Infrastructure: a trinity of software

CATALOGUE	REPOSITORY	PLATFORM
A place to share and discover research objects	Safely store and retrieve digital research objects	Analyse, simulate, visualize, process, manage, interact, ...
Unique (global) identifiers & versioning	Files (versioning, metadata)	Software services (web apps, APIs, Jupyter notebooks, workflow engines, VDIs)
Metadata (manual, automated) to facilitate discovery	Unique (global) identifiers & versioning	Compute (HPC/HTC), storage and networking
Actual object may be available (stored) elsewhere	Access policies (and sensitive data)	Collaboration and tiers
Catalogue and harvesting services	Long-term preservation	Generic-purpose and domain-specific
	File & Sync (personal space)	

A high-level overview of the Infrastructure's **architecture** (including Catalogue and Repository) is shown below Figure 17. The platform is implemented as a collection of microservices to deliver a Platform as a Service (PaaS) solution for managing DT models and executing simulation jobs.

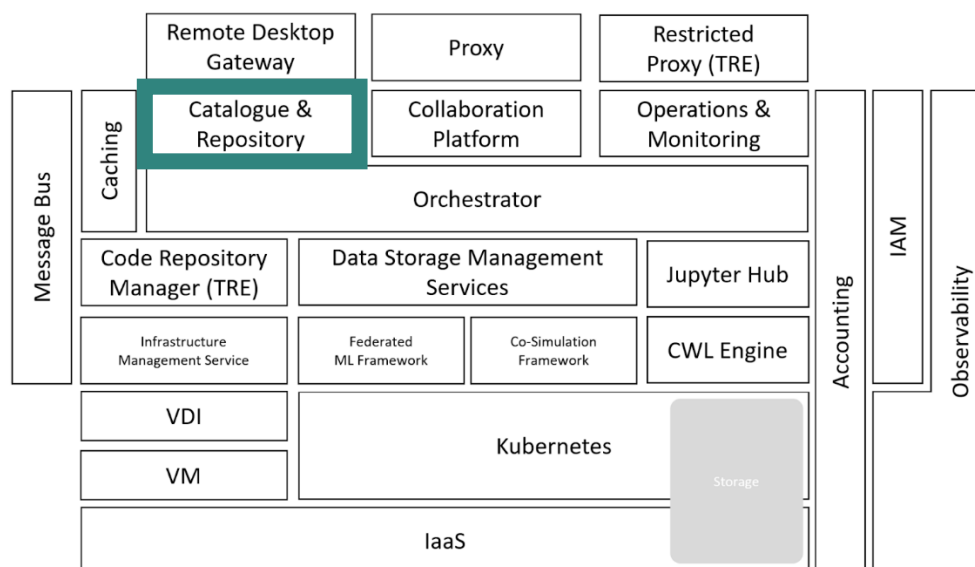


Figure 17: High-level overview of the proposed VHT infrastructure, including the Catalogue & Repository.

The component of **Catalogue & Repository** provides services for harvesting, storing, discovering and sharing digital models, datasets, notebooks and any artifacts (e.g., CWL workflows) related to digital

models and simulation. Resources published in the Catalogue require unique identifiers and are discoverable by third party services using established mechanisms. The **Catalogue federates VHT resources from other catalogues/repositories**, harmonizes any/all collected metadata, provides additional discovery facilities and modalities (*e.g.*, data profiling, visualizations) and access to the actual resources; these may be provided directly from suppliers (following any required authorization requirement) or from the repository (*e.g.*, when directly published in it from users or when replicated from a federated repository). Access to sensitive data takes place directly from the TRE, following a dedicated, traceable and safe data provision and use workflow.

16.4 Catalogue & Repository requirements

16.4.1 Requirements for both Catalogue and Repository

This section outlines the key functionalities and features of both the VHT Catalogue and Repository, emphasizing their intertwined roles in managing research objects.

- **Research Objects as first-class citizens:** the VHT Infrastructure treats Research Objects (RO), encompassing diverse resources such as data, models, notebooks, workflows, services, software, and simulations, as fundamental elements. The catalogue offers flexible viewing options to accommodate this diversity, including a unified catalogue or distinct instances tailored to different RO types.
- **Flexible and adaptable metadata schema:** a central requirement is a metadata schema that can describe the wide array of research objects within the VHT Infrastructure. The chosen approach should prioritize **flexibility and avoid imposing overly rigid standardization**. The system can employ a base schema like DataCite²¹² and introduce extensions to accommodate the specific attributes of different RO types. The schema should also support existing metadata schemata, enabling validation, storage, ingestion, transformation, and the ability to download metadata in various formats.
- **Robust system of identifiers:** unique and persistent identifiers are crucial for the management and retrieval of ROs. The VHT infrastructure will need to implement a system of **catalogue-specific unique stable identifiers** for ROs and users. Additionally, the system will automate the minting and assignment of Digital Object Identifiers (DOIs) to enhance the accessibility and recognition of ROs. Integration with external user identifiers like ORCID is a must.
- **Comprehensive file management capabilities:** the VHT's design should prioritize handling both data and models, accommodating the diverse needs of researchers. Files can be directly deposited or referenced, supporting various file types and offering options for validation, sanitization, transformation, and version control. The system should provide flexibility in managing file types, allowing users to work with open or closed vocabularies and supporting both raw and archived content.
- **Rich context through additional resources:** The repository should extend beyond the core ROs to include linked or deposited materials such as documentation, applications, and other ROs. This approach ensures that users have access to a **comprehensive collection of resources** that enhance their understanding of the research objects.
- **Streamlined publishing workflows:** the VHT Infrastructure should implement user-centric publishing workflows that allow registered users to contribute content. Before publication, content undergoes a review and vetting process by the Helpdesk to ensure quality and security. Organization-level control mechanisms enable entities like universities and laboratories to manage the publishing process. To enhance content quality, tiered curation processes for metadata, data, and models will need to be employed. An RO 'badge' system designates certain research objects as reference data.
- **Licensing, Access Policies, and RO Management:** The system supports a variety of **licensing options, including Creative Commons licenses and proprietary licenses**, to accommodate different levels of data sharing. Access policies range from public access to restricted access for confidential or sensitive content. To distinguish between ROs of varying quality and significance, the infrastructure can employ a **tiered curation process** including badges, and

²¹² DataCite Metadata Working Group. (2014). <http://doi.org/10.5438/0010>

crowd-sourced features like ratings and comments. The handling of RO derivatives, including capturing provenance and maintaining links to source ROs, is also considered.

- **User-friendly discovery tools:** discovering relevant ROs is crucial for the usability of the VHT Infrastructure. The system should offer various search and filtering capabilities, including plain-text search, advanced faceted search, thematic collections, and crowd-sourced favorites. Recommendations for ROs based on user behaviour, alternative content suggestions, and RO-specific user interface elements built on ontologies will enhance the discovery experience.
- **Enhanced functionality for improved usability:** The VHT Catalogue and Repository should be designed to go beyond basic storage and access by offering enhanced functionalities. These include RO-specific, intuitive viewers for easy visualization of content, automated data profiling to provide insights into data before downloading, integrated repository/platform actions to enable seamless integration with other tools and services, and linking of ROs with software and services for editing or running ROs.
- **Machine-discoverable content and metadata:** the VHT infrastructure should prioritize making content and metadata machine-discoverable, *e.g.*, by supporting Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH²¹³) for metadata exchange and interoperability. This will allow for easy integration with external repositories, platforms, and research infrastructures.
- **Comprehensive Statistics and Analytics:** the VHT infrastructure should collect and present statistics and analytics on content usage and impact. This data should be presented through public user dashboards displaying aggregated statistics, and private admin and publisher dashboards providing more detailed insights into user behaviour and content performance.

16.4.2 Requirements for the Repository

There are a number of requirements that only pertain to the VHT Repository, emphasizing storage, file management, and security considerations.

- **Multi-tier storage:** the VHT Repository should employ a multi-tier storage approach that utilizes different storage technologies based on content access patterns and sensitivity. This approach ensures efficient use of resources while addressing security and performance needs. Options range from high-performance file systems for active content to cloud storage for less frequently accessed data. The system also addresses specific requirements for managing sensitive data within TREs.
- **Enterprise File Sync and Sharing:** the VHT Repository should leverage Enterprise File Sync and Sharing (EFSS) to provide a robust and secure platform for collaboration. This approach enables seamless file sharing among individual users and collaborative groups, ensuring efficient data management and promoting teamwork within the EDITH ecosystem.
- **Extra services for data quality and security:** the repository should provide additional services that enhance data quality, accessibility, and security. These services include schema validation, data profiling, and data ingestion through various methods. The system should support anonymization and de-identification services to protect sensitive information, ensuring compliance with ethical and legal requirements.

16.4.3 Additional considerations

The VHT should implement **Single Sign-On** (SSO) to simplify user authentication and access to its Infrastructure. SSO will enhance the user experience by enabling users to access multiple related systems with a single login, streamlining their interactions with the platform.

Finally, the VHT Infrastructure needs to prioritize **interoperability with existing health data ecosystems and initiatives**, such as EHDS, EOSC and EBRAINS. The system will leverage international data standards to promote data sharing, reusability, and collaboration (*cfr.* PART 4). This interoperability ensures that the VHT seamlessly integrates with existing infrastructures, maximizing the impact and reach of the platform within the broader research community.

²¹³ <https://www.openarchives.org/pmh/>

17 Governance Principles

Corporate governance is founded on laws, policies, processes, systems and behaviours, and together they provide a system for the way in which an organisation is directed, administered and controlled. The governance framework sets out the roles, responsibilities and procedures for the effective and efficient conduct of its business. It is regularly reviewed so that it remains at the forefront of best practice. This chapter provides the principles for the VHT governance model.

17.1 Roles and Responsibilities

Well-defined user profiles are crucial for the success of the VHT Infrastructure, as they enable:

1. Efficient resource allocation;
2. Customized user experiences;
3. Robust data privacy and security;
4. Regulatory compliance;
5. Streamlined collaboration;
6. Improved monitoring and accountability;
7. Overall scalability and flexibility of the system.

In the VHT infrastructure, a **user profile** (defined as a collection of settings and information associated with a user) contains critical information that is used to identify an individual, such as their name, age and individual characteristics such as affiliations, background knowledge or expertise. The profile should be linked to an external IDP (IDentity Provider) and thus the authentication should be performed with trusted third party services, such as trusted National Identity Provider services in current EU states or an European wide citizens identification system such as the personal digital wallet²¹⁴ for EU citizens and residents, thanks to the trust framework created by the eIDAS Regulation in the long term. User profiles are adaptable as the individual changes jobs, relocates to other countries, or acquires other knowledge or expertise. The user profile does not constrain the specific roles of the user in the system. A **user role** is a well-defined collection of permissions within the systems allowing the user to achieve the role's intended purposes. In other words, user roles are clusters of system privileges that are designed to achieve specific goals using the VHT infrastructure. User roles are dynamic as their assignment to user profiles. A user profile can consist of various user roles as a person can have different aims in using the VHT platform in their capacity. New roles can be defined as new paradigms for using the VHT platform emerge.

User **roles** are a collection of capabilities that can be used to give access to concrete parts of the system. In the VHT infrastructure we can have the following roles:

- **Patient/citizens category**
 - Patient: access to their own VHT, personal medical data, and treatment options, enabling them to better understand their health and make informed decisions about their care.
 - Patient advocate: access to a specific patient's VHT and medical data with the patient's consent, allowing them to provide support and guidance in healthcare decision-making.
 - Citizen scientist: limited access to anonymized data sets and research tools to contribute to community-driven medical research initiatives.
- **Healthcare professional category (doctors, specialists, etc.)**
 - General Practitioner: access to their patients' DTs and relevant medical data to monitor health, diagnose conditions, and recommend treatments. Can create new simulations for their patients.
 - Medical specialist: access to group specific patients' DTs and related data within their area of expertise (*e.g.*, cardiologists accessing cardiac data) to provide specialized consultation and treatment recommendations.
 - Medical researcher: access to anonymized data sets and research tools to conduct studies and contribute to advancements in medical knowledge. Can run cohort analysis.

²¹⁴ https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_en

- Medical educator: access to VHTs for teaching and training purposes, allowing students and professionals to learn and practice in a simulated environment.
- **Creator/model developer category:** can upload new model versions, train the models.
 - Data scientist: access to anonymized data for developing models, as well as access to model training and evaluation tools.
 - Simulation engineer: access to simulation tools, environments (*e.g.*, cloud, HPC), and relevant data to design and validate virtual human twin simulations.
 - Model developer/owner: upload new model version, manage existing models (*e.g.*, configuration, model access constraints, *etc.*).
- **Platform administrator category**
 - System administrator: low-level access to the platform, along with a set of dedicated monitoring tools.
 - DevOps specialist: tasked with developing additional features in the VHT Infrastructure itself.
 - Data curator: tasked with curating the contributed resources.
 - Software developer: upload new model version, manage existing models (*e.g.*, configuration, model access constraints, *etc.*).

17.1.1 Role category: patient/citizen

The **patient/citizen** category encompasses end users who are primarily interested in the outcomes of medical therapy rather than in development of new treatment methods or direct processing of medical data. This category encompasses the roles of patient, patient advocate and citizen scientist.

Persons who belong to this category may be interested in using applications which encapsulate computational models **as long as appropriate user interfaces (UIs) are offered** by their respective developers. Patients/citizens may also want to submit their own personal data for processing, which implies the need for at least limited access to data storage and retrieval features. The ability to comment upon the features provided by the platform, and submit feedback could be regarded as “nice to have”, but is not regarded as essential for any prototype.

17.1.2 Role category: healthcare professional

A **healthcare professional** operates in a clinical setting and has a role in diagnosing and treating patients. This category encompasses the roles of general practitioner, medical specialist, medical researcher and medical educator.

Features provided to this category of users largely overlap with those provided to patients/citizens; however, additional access may be required, *e.g.* to enable processing of broader datasets. Unlike a citizen/patient, a healthcare professional should be able to review medical records of multiple patients (depending on their specific area of responsibility), and be able to perform cohort analyses and other studies which depend on aggregation of data representing multiple individuals. We assume healthcare professionals possess a low to moderate level of IT expertise; accordingly, the platform should expose dedicated GUIs for this group.

17.1.3 Role category: creator/Model developer

The **creator/model developer** category applies to experts who contribute computational artifacts and/or datasets to the platform. This category encompasses the roles of data scientist, simulation engineer and model developer.

Individuals who belong to this category are typically experienced users of IT platforms, including HPC and cloud computing systems, and may require lower-level access interfaces, or ways to interact with the platform/registry programmatically. Creators require guidelines or SOPs showing how to integrate existing computational models with the VHT Infrastructure, and the infrastructure itself must support such integration. We also believe it is essential for the platform and the repository to expose APIs, enabling deployment of higher-order tools or extension of any core UIs offered by VHT, as a step towards building an IT ecosystem.

17.1.4 Role category: platform administrator

The **platform administrator** category applies to users whose primary concern is to ensure continued operation of the EDITH platform. This category encompasses the roles of system administrator, DevOps specialist, data curator, software developer.

Administrators require low-level access to the platform, along with a set of dedicated monitoring tools. Software developers who belong to this category may be tasked with developing additional features in the EDITH platform itself - they should not be confused with model providers (who perform software development work on computational models). Administrators may work to expand the platform with additional hardware resources (computational/data storage), which requires bilateral contacts with their operators thereof. Legal aspects of such collaboration should be acknowledged, but are out of scope of this discussion.

17.2 Decision-Making Processes

17.2.1 Populating the VHT

As previously discussed, the data object is the central resource of the VHT, while the model objects operate in relating some input data objects and some output data objects. Here, we describe a possible approach to populating the VHT with data objects and model objects, graphically represented in the figure below.

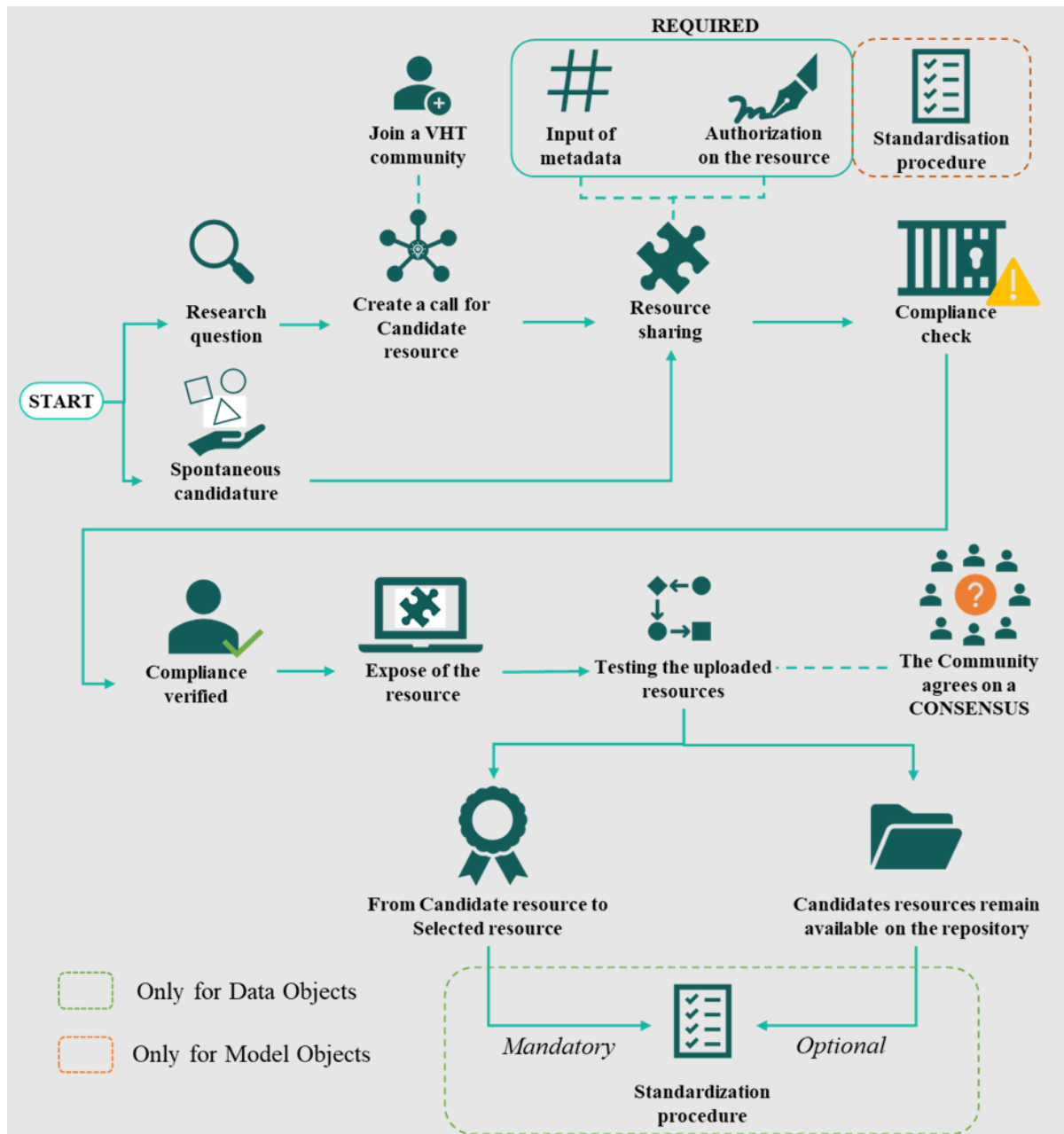


Figure 18: Procedure to populate the VHT

The process may begin starting from two different motivations:

- A potential user browses the catalogue in search of a specific data object, that is not present yet;
- A provider wants to share their data object and insert it in the VHT.

In the first case (a), the user creates a call for a candidate data object to respond to the research question. In this scenario, the user is advised to join the reference VHT community for that research question. In both cases (a and b), the providers start sharing the resources. Here, there are fundamental processes that must take place:

- Input of metadata:** the resource can be shared only if a minimum set of metadata is inserted, corresponding to the mandatory dataset defined in the VHT Metadata Schema, building on the vision elaborated in several EDITH deliverables and publications^{215,216,217};

²¹⁵ Deliverable D3.1: Vision for the VHT. <https://zenodo.org/records/7796845>

²¹⁶ Deliverable D3.2: first draft of the VHT roadmap. <https://zenodo.org/records/8200955>

²¹⁷ Viceconti et al., Position paper From the digital twins in healthcare to the Virtual Human Twin: a moon-shot project for digital health research. IEEE J Biomed Health Inform. 2023. <https://doi.org/10.1109/JBHI.2023.3323688>

- **Authorization on the resource:** the provider must prove to have the rights in sharing the resource;
- (Only for candidate model objects) **Standardisation procedure:** the candidate model object must be provided in a standardised format in terms of syntax and semantics.

For the last point, this is mandatory for candidate model objects since its absence would prevent its test and use within the VHT ecosystem.

Afterwards, the shared resource undergoes a **compliance check**, in such a way that the authorization presented by the provider is further assessed by a human operator. The resource is then displayed on the VHT catalogue, so that it is visible to visitors of the VHT ecosystem, together with the licence decided by the provider, which must be accepted by the user when using that resource.

The reference community for that resource **tests the resource**, giving a score in terms of credibility/validity/generalizability. Simultaneously, the reference community works towards a **consensus** on the topic referring to that resource if there are still some questions that need to be addressed, such as the ontology to be used. For a certain period of time, the resource will be tested for its integration in the VHT. There could be more than one resource responding for that candidate resource, or in general to represent a specific physiological mechanism. If testing is concluded successfully, then a resource will pass **from candidate resource to selected resource**, while the other resources will remain available in the repository to be used separately from the VHT. At this stage, if the resource is a dataset, the selected resource must undergo the **standardization procedure**, which is optional for those resources which remain in the “candidate” state.

17.2.2 Standardization procedure

Standardization is an essential requirement for the resources to be uploaded into the VHT Infrastructure to ensure interoperability. Here, standardization is referred to as giving a set of information related to the resource that helps, on one hand, in meeting the **FAIR principles**, and on the other to make VHT effective, since it will ensure a **sufficient level of interoperability**.

The application of standards imposes a rigid procedure to be applied. For this reason, in order not to discourage the providers and users to get involved into the VHT ecosystem, there are some steps that should be taken as optional. Here, we propose an approach to standardize resources of the VHT, displayed in the figure below. Based on the descriptors that exhaustively define a resource within the VHT (described in PART 2), we present the information contained in the standardization procedure and the challenges that still need to be addressed.

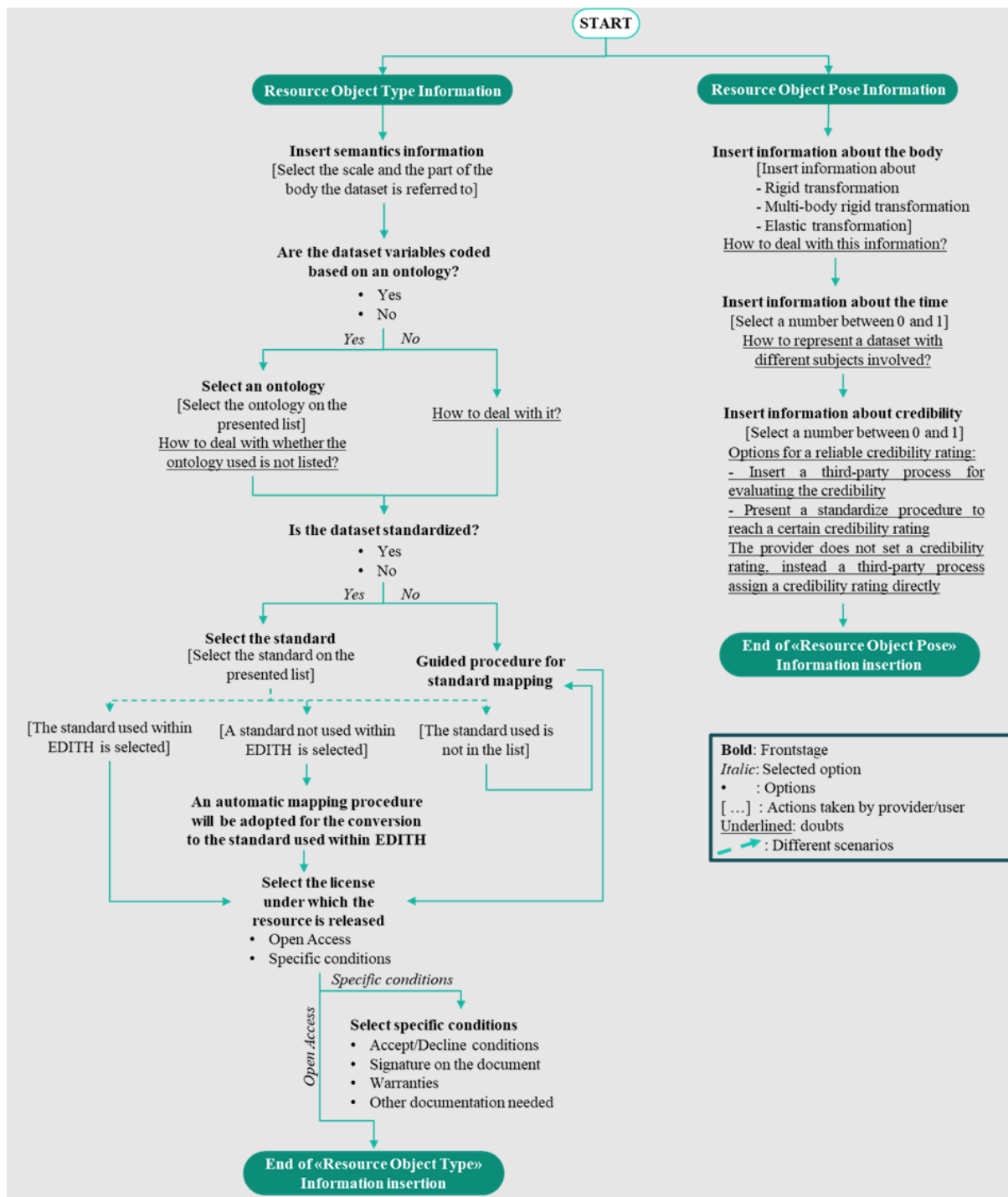


Figure 19: Standardization pipeline

The figure below shows the service blueprint representation of the ideal process of standardization for a dataset. Here, all the support processes needed to make the platform work are shown.

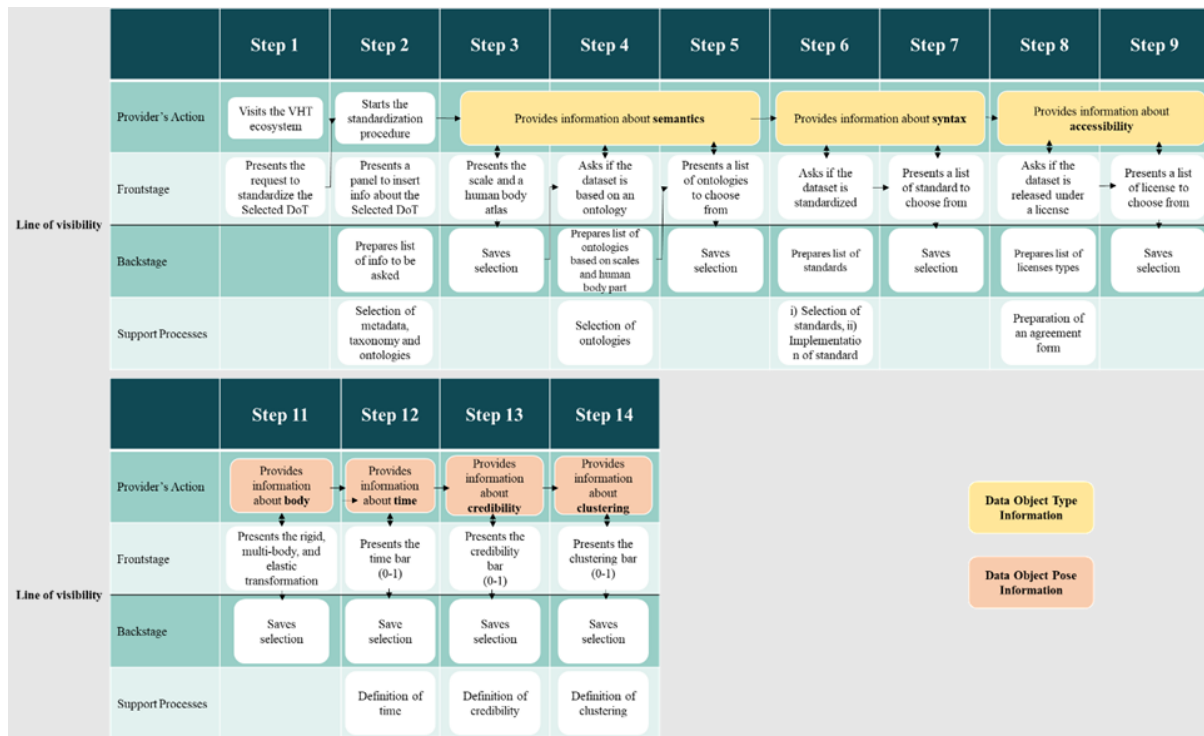


Figure 20: Service blueprint for data object standardization

17.3 Compliance and Legal Considerations

17.3.1 Premises and key issues

During the implementation of the VHT Infrastructure, the architecture outlined above is likely to change. Hence, the present section is not intended to cover every aspect of the governance of both the VHT Repository and its Catalogue. However, the main legal requirements are introduced in order to give a general and broad framework revolving around them. The legal considerations around the governance involve mainly three aspects:

- Data protection;
- Data sharing;
- Licenses for utilization of resources.

From the information collected up to now, it appears that third parties (“Provider(s)”) will be able to feed the Repository by uploading datasets, models, good practices and algorithms. It is important to consider that, in this stance, when the upload of the resources occurs, terms and conditions revolving around the licenses and the requirements related to data sharing and data protection should be immediately displayed and accepted by the provider. For example, one of the main requirements to be accepted for the uploading is agreeing that the resources will be visible to users of the Repository.

Defining the content of terms and conditions at the moment of the upload raises two different issues.

- 1) When the resource consists of a dataset, the liability in case of non-compliance is determined according to the existing norms in matters of data protection (General Data Protection Regulation, reg. 2016/679²¹⁸) and data sharing (Data Governance Act, reg. 2022/868²¹⁹). It seems therefore useful to dive specifically in the **allocation of liability** between providers and the VHT legal entity.
- 2) When the resource is a creative work, it is necessary to define if the work is released under a license and, if so, which kind of license. In this respect, solutions for the governance of the Repository may be various. Providers could be asked to define if their work is being uploaded under a **specific license**. However, this solution could present some criticalities in terms of

²¹⁸ <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

²¹⁹ <https://eur-lex.europa.eu/eli/reg/2022/868/oj>

management of the resources in the Repository. For this reason, it should be taken into account that the legal entity responsible for the VHT Infrastructure could propose a set of various licenses to all third parties, which guarantees the same rights of use of the uploaded resources to users.

A full analysis of all legal aspects related to the VHT in the broadest sense (not only its governance) is provided in PART 4.

17.3.2 Data protection

The resources present in the Repository and its Catalogue can contain both personal and non-personal data. In this regard, it seems useful to specify that the following considerations will apply only to datasets where personal information can be visible to users.

To the present state of the art, it is not possible to assess the role of the VHT legal entity in the management of the Repository and its Catalogue since solutions may vary according to the specific purposes and modalities of the data processing. Nonetheless, it seems that Providers and the VHT legal entity might have different liability regimes. Anyhow these aspects will be determined in the contract between the VHT legal entity and the Provider.

Considering the future development of the VHT's cloud-based federated Repository, its Catalogue and the modes of discovery, the paragraphs below will focus mainly on the general obligations defined by reg. 2016/679 for the licit processing of personal data by active subjects of processing.

Principles

According to GDPR (reg. 2016/679), data controllers must process data in compliance with the principles defined by Art. 5 GDPR. In particular, processing needs to be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed; the data controller must guarantee the update of data and their accuracy; personal data must be stored only for the necessary period and criteria for the definition of such period need to be provided. Overall, the data controller is responsible for compliance with the mentioned principles and is able to demonstrate it.

Legitimate Basis for Data Processing

Data can only be processed by the data controller where there is a legitimate basis under Art. 6 and 9 GDPR. Considering the purposes of scientific research (art. 9 lett. J), GDPR should be considered applicable for the processing occurring in the Repository and its catalogue. In this respect, one of the major criticalities to be taken into account consists in the fragmentation of conditions at the national level. According to Art. 89 GDPR, indeed, Member States can define more requirements for the legitimate processing of data for scientific research purposes.

Security Measures

Chapter IV, Section II of GDPR provides for the security measures that the data controller must adopt when processing personal data. Specifically, *“taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons”*, the controller needs to take appropriate technical and organizational measures to ensure an appropriate level of security. As a mere exemplification, the data controller can pseudonymize and encrypt personal data in a way that guarantees the appropriate safeguards for the data subject's rights. In assessing the appropriate level of security, account shall be taken in particular of the risks that are presented by processing, especially from accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to personal data processed.

Data Protection Impact Assessment

Considering the sensitive nature of data provided through the datasets, the data controller must conduct a Data Protection Impact Assessment (DPIA) according to Art. 36 GDPR. The DPIA must contain at least *“a systematic description of the envisaged processing operations and the purposes of the processing [...]; an assessment of the necessity and proportionality of the processing operations in relation to the purposes; an assessment of the risks to the rights and freedoms of data subjects”*.

Moreover, the data controller must provide for the measures adopted to avoid risks to the protection of personal data and the compliance with principles and obligations deriving from GDPR.

Privacy Policy

When processing data, the data controller is also obliged to provide to both the users and providers a privacy policy according to Art. 13 GDPR, defining the contents and limits of the pursued processing of personal data. It should define the identity and the contact details of the controller; the contact details of the DPO; the purposes of the processing for which the personal data are intended as well as the legal basis for the processing; the recipients or categories of recipients of the personal data; the fact that the controller may intend to transfer personal data to a third Country, and it is compliant with Chapter V GDPR. Moreover, the rights of the data subject should be highlighted as well as the period for which the personal data will be stored, or the criteria used to determine that period.

If personal data are being uploaded by a natural person different from the data subject, a legal person or a public entity, the contents of the privacy policy provided to data subjects need to be compliant with Art. 14 GDPR.

Active Subjects of Processing

The data controller can delegate the processing of personal data to data processors, according to Article 28 GDPR. The data processor provides “*sufficient guarantees to implement appropriate technical and organisational measures in such a manner that processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject*”. The processing conducted by the data processor needs to be described in a written contract or other legal act according to national or European law, that “*is binding on the processor with regard to the controller and that sets out the subject-matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller*”.

It is important to specify that, when the data processor does not comply with the conditions defined by the contract and, as a consequence, determines purposes and means of processing, they will be considered data controllers.

17.3.3 Data sharing

The Repository and its Catalogue are aimed at sharing resources with the users and among providers. Considering specifically the case of data being shared through datasets, the present paragraph is intended to focus mainly on the legal conditions set by the Data Governance Act (reg. 2022/868). However, it seems useful to premise that other European regulations may be applicable in the future due to the specificity of the data processed. In particular, the European Health Data Space (COM/2022/197 final²²⁰) will provide specific rules on sharing health data among data holders and data users for uses different from healthcare. Therefore, the following considerations are presented taking into account the developing legal framework at the European level.

Data Governance Act and General Data Protection Regulation

With specific regard to the purposes of the Repository and its Catalogue, Data Governance Act (DGA) aims at facilitating the sharing of data among data subjects and data users. In this respect, it is relevant to consider that DGA is not intended to derogate to provisions contained in GDPR. Consequently, personal data must be shared only in compliance with the data protection requirements, in particular in the presence of a legitimate basis

Data Altruism

DGA introduces the “data altruism” mechanism, consisting in the “*the voluntary sharing of data on the basis of the consent of data subjects to process personal data pertaining to them [...] without seeking or receiving a reward that goes beyond compensation related to the costs that they incur where they make their data available for objectives of general interest as provided for in national law [...] such as healthcare [...] or scientific research purposes in the general interest*”. Such a mechanism could be pursued by either recurring to data altruism organisations or using the European data altruism consent

²²⁰ <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52022PC0197>

form. In particular, when data is being shared directly by the data subject, consent could be collected through the form itself, as defined by Art. 25 reg. 2022/868. The form is made available by the European Commission in a uniform format, such that it should contribute to the consistent collection of consents across the Member States. It can be customised according to the specific sector of healthcare and prevention, and it guarantees that data subjects can give and withdraw consent from a specific data processing operation.

17.3.4 Licenses for utilization of resources

As due premises, it is relevant to consider that when providers make algorithms and software visible and possibly available, Dir. 2009/24/EC on the legal protection of computer programs is applicable. Except for the cases in which industrial law, with special regard to patents, needs to be taken into account, resources can be covered by copyright law. Providers will indeed be asked to specify the modalities in which economic rights related to their work will be transmitted through licenses. Hence, governance of the resources uploaded by providers could vary according to the level of permission to use their resources. In this respect, an alternative solution could consist of VHT providing the terms and conditions for a general license to be accepted by the providers in order to share their resources. It is relevant to consider that licenses can take on various forms according to the will of the parties. However, in the scientific literature, a general tripartition is offered in order to give a satisfactory overview of licenses:

- The **Proprietary Model**: it is based on the rights and conditions directly provided by the law. At the European level, the matter is regulated by both Dir. 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society and, more specifically, Dir. 2009/24/EC on the legal protection of computer programs. This license model guarantees that the holder of economic rights to the work has an exclusive right to use it in any form or manner for economic gain. The holder can authorize other people to use the work, but such authorization precludes any kind of modification or access to the source code, except for interoperability reasons.
- The **Open-Source Model**: generally, when the work is intended to be open source, it does not need to be protected by licenses. However, there may be specific elements that need to be guaranteed for the resource to be kept accessible and usable to the public. In this case, licenses could guarantee that the source code is accessible to the licensee, who can freely study it, modify it and share their own model with others, with the obligation of putting the same open-source license on the modified copies (copyleft clause).
- The **Open Access Model**: it allows the holder of economic rights to define more freely the conditions of use and modification of their work by licensees. A variety of clauses have been produced for the definition of the level of protection of economic rights according to this model. The holder of economic rights to the software can allow others to create copies and give credits to them; they can allow or restrict the possibility of accessing the source code and modifying the software; they can oblige the licensee to put the same kind of license on the modified copies. One general example of this category is the Creative Commons, which allows six different types of licenses according to the level of protection of the economic rights intended²²¹.

In any case, the user will have access to the resources provided by the third parties only when accepting the terms and conditions defined by the provider. The VHT Infrastructure will impose an obligation to display such conditions at the moment of the sharing.

²²¹ <https://creativecommons.org/share-your-work/cclicenses/>

18 Simulation platform – the basis

18.1 General considerations

The requirements and proposals listed in this chapter and the next are the results of a multi-faceted approach followed in the EDITH-CSA project. Initially, the **fundamental, universal requirements** gleaned from other research platforms were examined. These requirements were not tied to a specific research field, they were related to tools generally beneficial in the research community. Subsequently, the obtained considerations were aligned with prevailing **standards** in the VHT's areas of interest. Finally, pre-defined **EDITH use cases** (table below) as well as external contributions were examined to cater to the specific needs of researchers within particular fields within the VHT domain. This process refined both the problem statement and the ensuing solution, rendering them more tailored and specific.

Box 30: EDITH development – pre-selected use cases informing vision and roadmap

EDITH development: EDITH use cases

Developed by: BSC, U.Liège, UKA, QMUL, UNIBO and JFZ

From the beginning of the EDITH project, use cases pre-defined during the grant writing stage to be used as demonstrators have been studied from every angle (technology, infrastructure, ELSI, regulatory, market). The use cases had been selected to cover a wide range of characteristics in input, output, compute and workflow requirements. These use cases were complemented by externally contributed ones after a public call²²². All use cases were studied to extract requirements for the VHT infrastructure.

Table 2: Main characteristics of EDITH use cases

	Input	Output	Computational requirements	Resources within the UC
Cancer	<ul style="list-style-type: none"> • Molecular data • Clinical data • CT data 	<ul style="list-style-type: none"> • Image and video of the tumour 	<ul style="list-style-type: none"> • Local workstation • HPC facility 	<ul style="list-style-type: none"> • C++ scripts • R/Python scripts • Swift-T
Intensive care	<ul style="list-style-type: none"> • Patient characteristics • Real-time sensor data 	<ul style="list-style-type: none"> • Predictions on sensor data 	<ul style="list-style-type: none"> • Local workstation 	<ul style="list-style-type: none"> • Java scripts • Java containers
Cardio-vascular	<ul style="list-style-type: none"> • CT or MRI data • Electroanatomic mapping data 	<ul style="list-style-type: none"> • Simulation data 	<ul style="list-style-type: none"> • Local workstation • HPC facility 	<ul style="list-style-type: none"> • Docker for automatic segmentation • Cemrgapp • OpenEP • Meshtool • Paraview • OpenCARP • Python scripts
Osteoporosis	<ul style="list-style-type: none"> • Quantitative Computed Tomography • Patient data 	<ul style="list-style-type: none"> • Minimum Side Fall • Absolute Risk of Fracture at time 0 	<ul style="list-style-type: none"> • Local workstation • HPC facility 	<ul style="list-style-type: none"> • Python scripts • Ansys
Brain	<ul style="list-style-type: none"> • Cytoarchitectonic maps • Structural and functional connectivity • Cell densities • Receptor densities 	<ul style="list-style-type: none"> • Simulation data 	<ul style="list-style-type: none"> • Local workstation • HPC facility 	<ul style="list-style-type: none"> • Python scripts • Matlab scripts • TheVirtualBrain • Arbor

A simulation platform is regarded as complementary to the data repository discussed above. This layer comprises computational services deployed and accessible throughout the European ICT ecosystem. While referred to as a singular “platform” it can, in fact, comprise multiple autonomous services, which perform processing of data stored in the repository or retrieved through the catalogue. The platform – or its constituent services – should be able to interact with the previously described storage resources, and should enable deployment of computational models across organisational boundaries (figure

²²² <https://www.edith-csa.eu/call-for-use-cases/>

below). To this end, we recommend that VHT models should follow a uniform packaging scheme, such as provided *e.g.*, by containerization frameworks (Docker, Apptainer, etc.).

The added value of a VHT platform rests in the “three Rs”: Repeatability, Replicability, Reproducibility. Particularly in the context of medical simulations, the need to ensure validation and repeatability of computations, as well as to track their provenance, is particularly relevant, and any computational tool developed on top of the VHT data repository should enable such information to be tracked and accessed where necessary using widely-accepted packaging schemes.

The simulation platform should also expose dedicated interfaces to various groups of stakeholders, which correspond to the various user roles listed in the previous chapter. As described there, these interfaces (such as workflow management panels, data input forms, visualisation tools, and administrative mechanisms) can be implemented by autonomous tools, as long as standardised means of communication are maintained – by agreeing upon a set of shared APIs – and there is a single sign-on (SSO) as an authentication method for all the tools.

The VHT platform should analyse, simulate, visualise and process Research Objects, while, at the same time, should also facilitate the management of the results and the interaction between the users. As already mentioned, it will offer several software services (web apps, APIs, Jupyter notebooks, workflow engines, Virtual Desktop Infrastructures (VDI)) and will support several types of computational resources discussed next.

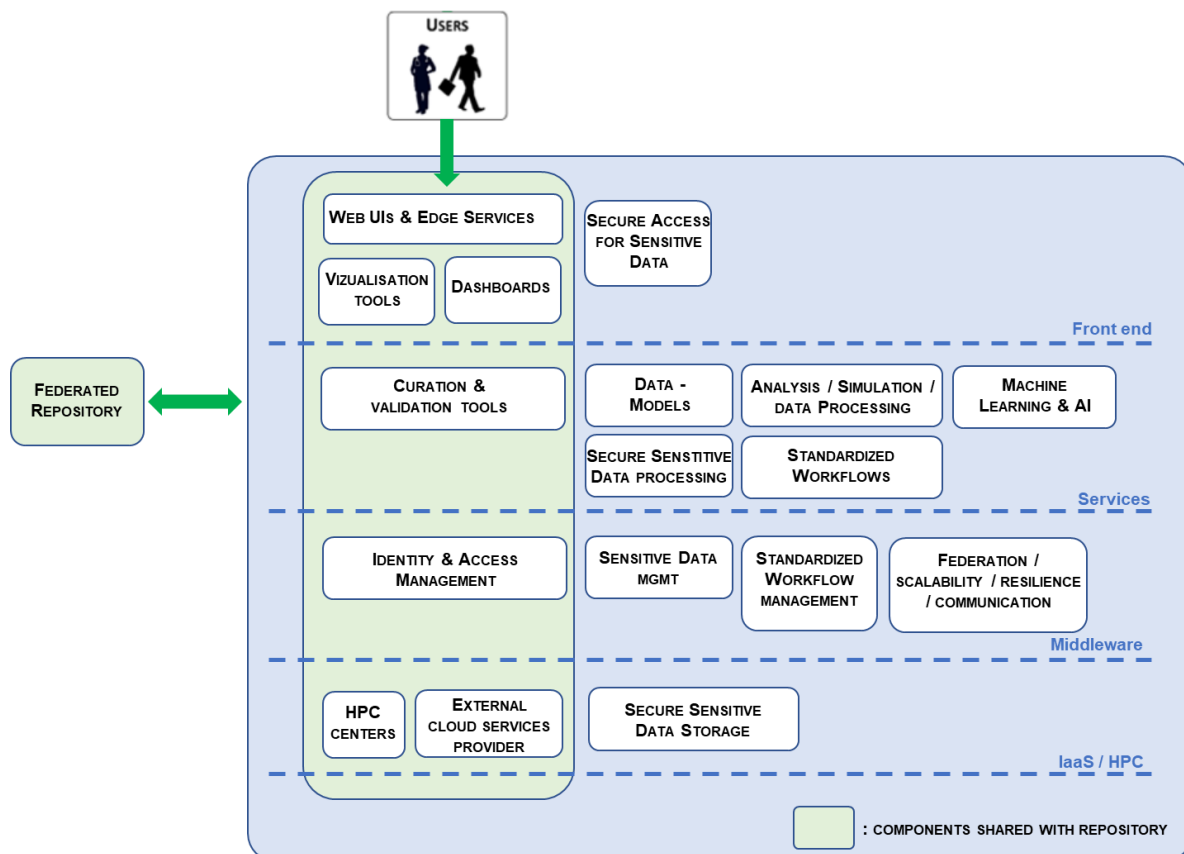


Figure 21: schematic representation of the simulation platform including the different layers and a (non-exhaustive) list of services.

18.2 General principles

The presence of a suite of loosely coupled and federated software and services offers **flexibility** and **scalability**, since individual components can operate independently, enabling easier updates, additions, and modifications without disrupting the entire platform. Moreover, a federated approach ensures **interoperability** between different software and services, especially if the platform needs to integrate with various tools and systems, providing users with a seamless and cohesive experience, and

adaptability, since changes or updates to one component do not necessarily affect the others, promoting adaptability to evolving technologies and user requirements.

The VHT platform should therefore:

- *be **expandable by design (Core Building Blocks)***. A platform designed for expansion allows it to grow and evolve over time. As user needs change or new functionalities are required, the platform can easily integrate additional features without major overhauls. Core building blocks enable a modular approach to development. This modular structure facilitates the addition of new features or the replacement of existing ones without disrupting the entire system. Moreover, developers can focus on enhancing specific components, or building new ones, without needing to understand the entire platform.
- ***process, analyse, visualize, simulate***. The ability to process, analyse, visualize, and simulate data covers a wide spectrum of research and analytical needs. Users can perform end-to-end workflows within the platform, from data input to visualization of results.
- *have platform-wide **access policies, tiers, and quota***.

Access policies ensure that users have appropriate permissions based on their roles and responsibilities. This is crucial for maintaining the security and integrity of the platform, while tiers and quota provide a structured way to allocate resources based on user needs. This ensures fair usage, prevents misuse, and optimizes resource utilization across the platform.

18.3 VHT platform user profiles and user roles

The VHT platform access and data management needs to be realised through an adaptable, secure, and efficient system, equipped to answer the dynamics of healthcare across EU borders, utilising existing data, models and state-of-the-art computational resources.

While designing and implementing user profiles, user roles, and their capabilities within the VHT platform (described in the previous chapter), the following considerations are important.

- **Efficient allocation of resources**: understanding the specific needs and usage patterns of different roles allows to allocate resources, such as computational power and storage, more efficiently. The platform should aim to harmonise user needs with system capacity, ensuring optimal performance, thereby contributing to cost-effectiveness and environmental sustainability.
- **Customised User Experience (UX)**: knowing the unique needs of different user roles, the platform should be designed to deliver a tailored user experience. By offering the most relevant features and data in a user-friendly interface, productivity, user satisfaction, and ultimately, patient outcomes can be boosted. For example a clinician-researcher might only have access to resources for which there is a minimal quality threshold met, whereas academic researchers might see the entire range of models regardless of their quality/credibility.
- **Data privacy and security**: in the era of data breaches, the need for robust data privacy and security mechanisms cannot be overstated. Well-defined user profiles are the foundation for implementing stringent access control, ensuring that each user group only accesses data and functionality relevant to their roles, thereby safeguarding sensitive information.
- **Compliance with regulations**: healthcare data is governed by strict privacy and security regulations. Well-defined user profiles and user roles enable the platform to enforce compliance effectively, ensuring user access aligns with legal requirements, such as GDPR in Europe (*cfr.* previous chapter).
- **Streamlined collaboration**: by facilitating clear distinctions and interactions among user roles, the platform should streamline collaboration among healthcare professionals, researchers, and regulatory agencies. This is a step towards a future of improved patient outcomes and accelerated medical advancements, driven by efficient information exchange and mutual growth.
- **Monitoring and accountability**: clear user profiles form the basis of effective monitoring, aiding in tracking user activity and identifying potential areas for improvement. This accountability is essential for system integrity, encouraging responsible usage, and enabling swift responses to unauthorised or suspicious activities.

- **Scalability and flexibility:** as the scope of the VHT expands, well-defined user profiles will ensure the platform's ability to scale efficiently and adapt to new user groups and requirements. This flexibility is key to the platform's long-term viability, as it allows us to evolve alongside the ever-evolving healthcare landscape.

Recognizing the diversity and complexity inherent in the stakeholders backgrounds, experience and intended use, well-defined user profiles and roles will play an important role in system efficiency, data security, regulatory compliance and user experience.

18.4 The technological starting point

18.4.1 Scientific and computational workflows and workflow engines

One of the most important and crucial explorations lies in the realization, implementation, and execution of computational workflows within the VHT Infrastructure. In the domain of Life Sciences, the tied description of scientific and computational workflows with the system that is responsible to execute steps in a workflow, handle failures and monitor workflows belongs already to the past^{223,224,225}. Many **workflow management systems**²²⁶ have shifted their roadmaps to support **integration with open, common, standard frameworks**. After a careful consideration of various ways of defining workflows in a structured, open and standardised way, using and embracing evolving technologies and tools which grow over time, **Common Workflow Language (CWL)**²²⁷ standard is recommended as a key component in the realization of the VHT Platform. CWL offers an open standard syntax for describing tools that are used as steps for data analysis, visualization and simulation workflows. CWL's syntax is compatible with a plethora of workflow engines and workflow management systems²²⁸ facilitating execution, handling errors and reruns in the event of failures. CWL workflows are strongly connected with metadata fetched during the runtime (retrospective) as well as while they are authored (prospective) utilizing the Research Object Crate (RO-Crate)²²⁹ initiative. This initiative aims at lightweight packaging of research data with metadata and, thereby, adds value concerning the trinity of the one-stop shop for the VHT Infrastructure, enhancing the overall functionality and utility of the Infrastructure.

In addition to the workflow language, one or more **workflow engines** need to be selected to serve as core VHT Infrastructure workflow management system service. The recommendation is to manage computational workflows on top of a variety of underlying infrastructure, spanning from cloud-based systems like OpenStack²³⁰, to HPC^{231,232,233} for high intense workloads, or container-based solutions such as Kubernetes²³⁴. These workflow management systems will be responsible to execute, oversee monitoring of workflow steps, rerun steps or whole workflows in the event of failures, fetch data to where the execution takes place and more. Finally, as the CWL **community** continues to expand and grow, refine the specification and enhance tools and technologies, support for features like Graphics Processing Units (GPU), Message Passing Interfaces (MPI) and integration with systems such as SLURM²³⁵ or LUSTRE²³⁶ has become increasingly viable and standardised. Most of the workflow engines **compatible with HPC systems** have been developed or improved to efficiently handle these features. Furthermore, community-driven efforts have established best practises for seamlessly integrating these features into CWL workflows, simplifying researchers' access to HPC resources while

²²³ https://doi.org/10.1162/dint_a_00033

²²⁴ <https://doi.org/10.1371/journal.pcbi.1008622>

²²⁵ <https://arxiv.org/pdf/2105.07028.pdf>

²²⁶ <https://github.com/common-workflow-language/common-workflow-language/wiki/Existing-Workflow-systems>

²²⁷ <https://doi.org/10.1145/3486897>

²²⁸ <https://www.commonwl.org/implementations/>

²²⁹ <https://doi.org/10.48550/arXiv.2108.06503>

²³⁰ <https://www.openstack.org/>

²³¹ <https://www.ibm.com/topics/hpc>

²³² <https://www.intel.com/content/www/us/en/high-performance-computing/what-is-hpc.html>

²³³ <https://www.nvidia.com/en-us/glossary/high-performance-computing/>

²³⁴ <https://kubernetes.io/>

²³⁵ <https://slurm.schedmd.com/documentation.html>

²³⁶ <https://www.lustre.org/>

leveraging CWL for workflow description and execution. Some noteworthy currently reviewed examples, that adhere to these requirements are Streamflow²³⁷ and Arvados²³⁸.

18.4.2 Jupyter notebooks

Computational workflows are not the only key component in the realization of the VHT Platform. The evolution of Jupyter notebooks²³⁹ and the noteworthy innovations like JupyterLab make them a valuable component for the VHT Platform as well. The **enhanced user experience**, offering a more unified workspace for users to select from various parameters, including kernels (experimental or production ready ones), memory allocations and access to GPUs, to create their own workstations is of great importance for VHT users. Another significant aspect of Jupyter Notebooks involves the potential **integration of interactive notebooks with HPC systems**. This integration facilitates the efficient handling of high-intensive workloads, empowering the execution of tasks, simulations and data analyses by leveraging the parallel capabilities within HPC systems via Unicore system, in cases where interactivity must be in place and computational workflow standardization is not the primary goal.

An additional advantage of Jupyter notebooks is their capability to **host on-demand ephemeral services** within the VHT Platform. This feature grants users the convenience of running live interactive code seamlessly – possibly – integrated with HPC components. By incorporating lessons learned and successes achieved by other Research Infrastructures, such as EBRAINS²⁴⁰, by employing dedicated build and delivery processes for JupyterLab environments, the VHT can position itself in the forefront of realizing initiatives that have proven efficient in large consortia. The mentioned implementation centralizes a variety of tools and services, spanning from analyses, visualization and simulation of micro- to macro- to whole brain scale, in a unified repository and a structured system with kernel releases and version controls. This initiative was implemented via the Spack manager system²⁴¹, enabling users to create pipelines and interactive workloads for Jupyter Notebooks. In that way, scientists meet their specific scientific requisites to foster innovation in a large variety of (neuro-) scientific areas.

18.4.3 Infrastructure resource usage cost

Envisioning the future with the emergence of a one-stop-shop VHT Infrastructure, integrating an **accounting component** becomes important. The responsibility involves overseeing the VHT Platform's resource usage costs^{242,243,244} by VHT users. To prepare this, a monitoring phase is first required, of which the primary objective is to enhance observability and forecast future needs based on computing, storage and network allocations, optimizing resources usage and conducting risk assessments. This is significant for the VHT Platform's expansion, enabling the tracking of demand while seeking, accessing, managing and modifying Research Objects – resources such as data, tools, workflows – within the VHT Infrastructure. In general, understanding the **compensation of resources** is crucial in business analytics for a) budgeting, b) cost control, c) optimizing resource utilization. This process includes evaluating the expenses associated with acquiring, maintaining and utilizing resources to achieve maximum output with minimal input, ensuring efficient resource allocation, avoiding underutilization or over-provisioning, and continuously seeking opportunities for optimization.

For example, Public Health Scotland's National Safe Haven²⁴⁵ is a secure research environment that allows researchers access to health-related data while maintaining strict privacy and security measures. Utilizing cloud services, especially with elastic features, enables the Safe Haven to adapt its computational resources according to researchers' needs, ensuring efficient and secure data processing while controlling costs. Cloud computing services like AWS, Azure and private ones, provide flexible, scalable, and cost-effective solutions. While allowing users to maintain robust security and compliance measures.

²³⁷ <https://doi.org/10.48550/arXiv.2002.01558>

²³⁸ <https://doi.org/10.5281/zenodo.6382942>

²³⁹ <https://jupyter.org/>

²⁴⁰ <https://www.ebrains.eu/>

²⁴¹ <https://spack.io/>

²⁴² <https://www.finops.org/landscape/>

²⁴³ <https://www.opencost.io/>

²⁴⁴ <https://www.kubecost.com/>

²⁴⁵ <https://www.isdscotland.org/products-and-services/edris/use-of-the-national-safe-haven/>

18.4.4 Authorization and accounting

Integration with Enterprise File Sync and Share (EFSS) systems for secure storage of Research Objects as well as personalized dashboards for execution and/or customization of Research Objects are additional crucial elements for the realization of the VHT Infrastructure. Both initiatives take collaboration between various users one step further by categorizing them into distinct tiers, granting varying levels of access and varying levels of actions that can be made within an organization. For the VHT Infrastructure specifically, having the possibility to **categorize VHT users** to grant access to various data repositories is of great importance. In general, users undergo authentication processes^{246,247,248} to verify their identities, subsequently **gaining authorized access** to designated resources based on their roles. Additionally, EFSS systems often **allocate credits or quotas** to users, restricting usage beyond predefined limits. This credit system promotes responsible data handling and ensures efficient resource allocation. This strongly relates to the accounting system mentioned in the previous section. Implementing Identity and Access Management (IAM) strategies within EFSS reinforces security protocols by enforcing policies across services, regulating user access and monitoring activities. These measures encompass advanced authentication mechanisms and role-based access controls, fostering a **secure collaborative environment**.

As an example, the Beacon project²⁴⁹ pioneers standardised and secure data sharing practices across diverse institutions in the scope of the GA4GH initiative. Leveraging advanced security protocols, the Beacon project incorporates cutting-edge authentication and authorization methods, aligning with security standards. The project's achievements in security, interoperability and access control methodologies serve as blueprint, influencing the evolution of EFSS systems by setting new benchmarks for secure data handling and collaborative research endeavours. GA4GH Passports and the Authorization & Authentication Infrastructure²⁵⁰ is also an initiative that is worth mentioning. from other research activities and external threats.

18.4.5 Trusted research environments and the 'five Safe principles'

Trusted Research Environments (TREs)²⁵¹ and the 'five (5) Safe principles'²⁵² were discussed earlier in this PART 4 of the roadmap, for their added value for exploration within the VHT Platform's implementation. Here, some crucial 'Safe Principles' are elaborated, focusing on 'Safe Settings', 'Safe Outputs', and 'Safe Computing'^{253,254}, which are strongly associated with the VHT Infrastructure.

Beginning with 'Safe Settings' and 'Safe Computing', solutions offer secure technical approaches for analyses, storage and access across various infrastructure layers. One of the critical approaches is the utilization of **Virtual Desktop Infrastructures** (VDIs), offering flexibility and ephemeral solutions, particularly beneficial in scenarios involving sensitive data that must remain on-premises or high-intensity workloads causing substantial delays in bringing/moving data to where execution takes place. The VHT Platform gains additional value by offering VDI solutions, including customized virtual machines for desktop workloads and applications, aligning with the paradigm of numerous TREs in non-EU countries such as National Institutes of Health (NIH), SAIL – Secure Anonymised databank²⁵⁵ and OpenSafely²⁵⁶. As seen in TREs such as OpenSafely and Genomic England²⁵⁷, users are afforded access to interfaces facilitating the creation and execution of Machine Learning algorithms. These platforms support the development and execution of workflows²⁵⁸ scripted in Nextflow and Workflow Descriptive Language (WDL), operationalized through the Cromwell workflow engine. In addition, Nextflow demonstrates versatility by containerizing R, Python3, bcftools, and VEP (Variant Effect

²⁴⁶ Authentication and Authorization Infrastructure (AAI), <https://data.europa.eu/doi/10.2777/8702>

²⁴⁷ <https://aai.grnet.gr/>

²⁴⁸ <https://www.keycloak.org/>

²⁴⁹ <https://doi.org/10.1093/bioinformatics/btac568>

²⁵⁰ https://www.ga4gh.org/news_item/ga4gh-passports-and-the-authorization-and-authentication-infrastructure/

²⁵¹ <https://doi.org/10.2196/33720>.

²⁵² <https://doi.org/10.5281/zenodo.5766512>

²⁵³ <https://doi.org/10.5281/zenodo.5766512>

²⁵⁴ <https://doi.org/10.5281/zenodo.4594703>

²⁵⁵ https://saildatabank.com/wp-content/uploads/2023/11/SAIL-Annual-Report-2022-23_AMENDS.pdf

²⁵⁶ <https://doi.org/10.3399/BJGP.2022.0301>

²⁵⁷ <https://www.genomicsengland.co.uk/>

²⁵⁸ https://re-docs.genomicsengland.co.uk/small_variant/, https://re-docs.genomicsengland.co.uk/sv_cnv/, https://re-docs.genomicsengland.co.uk/sv_cnv1/

Predictor). While WDL and CWL share similarities, it is noteworthy that CWL encompasses important features that WDL lacks²⁵⁹, offering enhanced functionalities in the proposed VHT Platform implementation.

Another effort strongly relates to ‘Safe Settings’ that could have a potential value in the context of the VHT Infrastructure and particularly the VHT Platform, is the **support of a curated selection of package libraries for often-used technologies** in the medical care areas and personalized healthcare such as R, Python Conda, SAS, STATA, and SPSS, ensuring security through a whitelist in code and library import. In some TRE environments^{260,261}, specialized technical teams can accommodate requests for the inclusion of extra packages, if this is needed, after performing a strict evaluation. Furthermore, these systems integrate robust security measures, including AV/Malware protection, SEIM (Security Information and Event Management), and Network Monitoring to fortify the underlying infrastructures. In terms of **collaboration software**, some TRE platforms incorporate Git for version control, Wiki and Confluence for documentation and information sharing, along with Shared File Store and Shared Databases for collaborative projects. Moreover, TRE environments²⁶² enable federated queries that allow users to access distributed data sources without moving the results off-premises, ensuring data confidentiality and compliance with security protocols. All initiatives are important to the one-shop stop of the VHT and could gain their own component implementation in the VHT Platform.

Finally, a closer look to ‘Safe Principle’ that concerns the ‘**Safe Output**’ became important. Within some TREs such as Secure eResearch Platform (SeRP)²⁶³, Safe Output means that **no actual data can be moved outside the premise** once the analyses, simulation, workflows have been executed. But there are some whitelisted and aggregated results that users may extract from. Specifically, SeRP facilitates the creation of aggregated graphs and tables following either rule-based or principle-based methodologies. To fortify security, Airlock²⁶⁴ is predominantly employed, allowing whitelist-based control over exported data, ensuring strict compliance measures within the system. All activities regarding ‘Safe Output’ must undergo human review and **adhere to established governance protocols**. In some cases, exporting Health Electronic Records (HER), images, and genomics data is feasible post-human review, ensuring compliance and data integrity. Last but not least, data transmission to the Biobank and utilization of the Scottish National Safe Haven are important initiatives to take into consideration for the VHT Platform. The latter one leverages GA4GH's Beacon Application Programming Interface (API) to enable federated queries, supporting distributed data access.

18.4.6 Computational infrastructure

The development of research infrastructures occurs against the background of ongoing emergence of exascale-level systems, coupled with new management and (hardware/software) architectural approaches, focusing on modular, integrated and lightweight solutions, which enable end-users across all scientific domains to use these future **supercomputing systems**²⁶⁵. While most VHT models can be simulated on standard setups, some scenarios require considerable computing power to achieve the desired accuracy. In these cases, future exascale HPC systems (up to 10^{18} FLOPS) will enable the development of high-resolution VHTs that model the dynamic behaviour of human tissues and organs across multiple spatial and temporal scales. As the number of entities being modelled – such as cells and proteins – grows, complexity increases exponentially, making it feasible only by leveraging the combined power of multiple interconnected machines in HPC facilities.

In parallel to this, as discussed in PART 2, advanced hardware architectures and hardware-software co-design such as **cloud/fog/edge computing** are essential to optimize performance, reduce power consumption, and ensure scalability for personalized healthcare applications in the VHT.

There is also ongoing work on addressing organisational and legislative issues involved in sharing and processing scientific data in federated environments, as evidenced *e.g.*, by the European Health

²⁵⁹ <https://doi.org/10.1038/s41598-021-99288-8>

²⁶⁰ <https://www.isdscotland.org/Products-and-services/Edris/>

²⁶¹ <https://serp.ac.uk/about-serp/>

²⁶² <https://trefx.uk/>

²⁶³ <https://serp.ac.uk/about-serp/>

²⁶⁴ <https://www.airlockdigital.com/airlock-application-whitelisting-features>

²⁶⁵ https://doi.org/10.31577/cai_2020_4_617

Research and Innovation Cloud²⁶⁶. Below, a non-exhaustive overview is provided, highlighting some initiatives being highly relevant to the computational aspects of the VHT Infrastructure. Synergistic and strategic collaborations with these initiatives will allow to identify more effectively the unmet needs and services complementing those offered elsewhere.

A key initiative on advanced computing and data management is **PRACE**²⁶⁷. This bottom-up science-driven initiative complements the top-down initiative of **EuroHPC**²⁶⁸, which supports the development of supercomputing systems across Europe. In addition, many national and regional resource federations, such as the **PL-Grid** infrastructure²⁶⁹, augment and extend the European-level efforts, providing additional opportunities related to academic studies and research data processing.

In this HPC context there are several centres of excellence focused on biomedical applications. **CompBioMed**²⁷⁰ is a Centre of Excellence focused on the use and development of computational methods for biomedical applications. **PerMedCoE**²⁷¹ is an HPC/Exascale Centre of Excellence for Personalised Medicine in Europe. For instance, PerMedCoE's tools observatory and community benchmark will provide interesting input when discussing the multiscale integration of resources and the model validation in the creation of the VHT applications. **MICROCARD**²⁷², is a Centre of Excellence developing a DT of the cardiac muscle at the micrometre scale together with tailored simulation software and workflows

The **Fenix** research infrastructure²⁷³ has emerged as a major provider of **federated supercomputing and storage resources**. The distinguishing characteristic of the Fenix e-infrastructure is that data repositories and scalable supercomputing systems are in close proximity and well-integrated. Fenix has been developed in as part of the Human Brain Project (HBP) to enable computation and data intensive research. **GEANT**²⁷⁴ interconnects Europe's national research and education networks with a high - bandwidth, high-speed and highly resilient pan-European network. **EGI**²⁷⁵ is an international federation delivering open solutions for advanced computing and data analytics in research and innovation.

In 2024, the European Commission introduced a comprehensive package of measures aimed to establish Europe as a global leader in AI innovation, by supporting startups and SMEs in developing trustworthy AI aligned with EU values and regulations. One element of this package are the **AI Factories**²⁷⁶, envisioned as collaborative ecosystems integrating computing power, data, and expertise to advance generative AI. Launched in 2025, they aim to drive innovation across sectors such as healthcare, manufacturing, climate, and finance by fostering cross-border collaboration among universities, supercomputing centres, industries, and financial institutions. These AI Factories will deploy advanced AI-optimized supercomputers and upgrade existing systems, significantly enhancing Europe's AI capabilities. This initiative aligns with the broader EU strategy to create an ecosystem of excellence and trust in AI.

18.5 Collaborations with other platforms

Today, repositories and platforms come in many shapes and sizes; there is a large number of **Research Infrastructures** (RIs) and the number is growing fast due to the role of science as a driver of innovation and economic growth, the ever-increasing importance of information technology in science, and the benefits that accrue from greater collaboration and scale. There are a lot of RIs that support medical sciences. The vast majority²⁷⁷ are “single-site” (small range of services, limited number of users), while about only one third of the RIs listed are “distributed”, with facilities spread across multiple locations.

²⁶⁶ <https://doi.org/10.1186/s13073-020-0713-z>

²⁶⁷ <https://prace-ri.eu/>

²⁶⁸ EuroHPC portal, <https://eurohpc-ju.europa.eu/>

²⁶⁹ PL-Grid infrastructure, <https://www.plgrid.pl/en>

²⁷⁰ <https://www.compbioimed.eu/>

²⁷¹ <https://www.permedcoe.eu/>

²⁷² <https://microcard.eu/index-en.html>

²⁷³ <https://fenix-ri.eu/about-fenix>

²⁷⁴ <https://geant.org/>

²⁷⁵ <https://www.egi.eu/>

²⁷⁶ <https://digital-strategy.ec.europa.eu/en/policies/ai-factories>

²⁷⁷ <https://portal.meril.eu/meril/>

In the 2021 roadmap²⁷⁸ of the European Strategy Forum on Research Infrastructures (ESFRI), there are 16 RIs in the Health and Food category and four in Data, Computing and Digital Research Infrastructures, all 20 of them of distributed type. For example, **ELIXIR**²⁷⁹, is a sustainable distributed European infrastructure that brings together life science resources where scientists can access biological data, software, cloud, storage and supercomputers from a single infrastructure. **BBMRI-ERIC**²⁸⁰ brings together all the main players from the biobanking field – researchers, biobankers, industry, and patients – to boost biomedical research. They offer quality management services, support with ethical, legal and societal issues, and a number of online tools and software solutions. As a last example, **EATRIS**²⁸¹, the European infrastructure for translational medicine brings together resources and services for research communities to translate scientific discoveries into benefits for patients.

The VHT should actively collaborate with initiatives inside the EU, focusing on sharing, accessing data and services across the pan-European layer, in a lawful, ethical manner via legal frameworks for various scientific areas. In this section we focus on three relevant initiatives that are important from a technical point of view, namely **EBRAINS RI**, the **European Open Science Cloud (EOSC)** and the **European Health Data Space (EHDS)**.

Considering recent cutting-edge solutions and achievements made by the EBRAINS Platform, one worth mentioning is the seamless integration of various components, spanning from simulation engines and software tools to the EBRAINS platform itself. The ultimate objective was to grant users access to a wide array of analysis, simulation and visualization tools without requiring any actions related to installation, loading or combining them to create pipelines and workflows. In light of these achievements, Standardised Workflows are the first-class citizens at EBRAINS RI. Within this context, workflows are described using a common, open and standard way via Common Workflow Language. These workflows have unique Digital Object Identifiers (DOI), allowing users to find, discover, access and request their addition to **EBRAINS Knowledge Graph (KG)**²⁸². They are associated with input and output data and models which are used during runtime, incorporating significant prospective and retrospective metadata. This comprehensive description and execution methodology expands innovations and interoperability. These standardised workflows are executed and monitored in the EBRAINS underlying infrastructure via workflow engines and workflow management systems. The interconnection of the VHT Platform and Catalogue with the EBRAINS KG holds substantial importance.

Regarding the **European Open Science Cloud (EOSC)**²⁸³ initiative; it stands as a transformative initiative fostering open science by providing a unified platform for accessing, sharing, and utilizing research data and services across disciplines. At its core, EOSC encompasses key facets like accessibility, interoperability, sustainability, and collaboration, aiming to facilitate transparent and reproducible research practices. Integral to this ecosystem is the EOSC Marketplace²⁸⁴, serving as a central hub where researchers can discover, access, and integrate various digital resources and services seamlessly. This marketplace streamlines resource discovery, promotes interoperability, and encourages community engagement, aligning with EOSC's mission of advancing open science. The commitment to future readiness of the VHT Infrastructure should include plans to support metadata schemata focused on RIs (such as DCAT, DCAT-AP, and CERIF), and to be informed by technical interfaces and APIs aligned with the EOSC Interoperability Framework (EIF). These endeavours ensure that the proposed VHT metadata schema remains aligned with broader global research data and infrastructure efforts, fostering effective communication, integration among repositories, and adherence to EOSC's overarching goals.

Finally, the **European Health Data Space (EHDS)**²⁸⁵ is a strategic initiative by the EU that aims to establish a secure and interconnected platform for the exchange and utilization of healthcare data across EU member states. Its goal is to facilitate seamless access to health data while upholding strict data

²⁷⁸ <https://www.esfri.eu/esfri-roadmap-2021>

²⁷⁹ <https://elixir-europe.org/>

²⁸⁰ <http://www.bbmri-eric.eu>

²⁸¹ <http://eatris.eu>

²⁸² <https://search.kg.ebrains.eu/>

²⁸³ <https://eosc-portal.eu/>

²⁸⁴ <https://marketplace.eosc-portal.eu/>

²⁸⁵ <https://www.european-health-data-space.com/>

protection standards, fostering interoperability among healthcare systems. The collaboration between the VHT Platform and the EHDS lies in the fact that VHT relies on health data as a fundamental source of its realization, development and implementation. Defining the essential issue of primary and secondary use of health data, makes EHDS of pivotal importance for VHT technology to move forward and beyond.

Box 31: Success story – SEEK platform for heterogeneous scientific research outputs

Success story: a resource designed to facilitate the management, sharing, and exploration of heterogeneous scientific research outputs

Status: publicly available

Website: <https://seek4science.org/>

The SEEK platform, developed by FAIRDOM, is an open-source, web-based resource designed to facilitate the management, sharing, and exploration of heterogeneous scientific research outputs, including datasets, models, simulations, processes, and results. It emphasizes the preservation of associations among these outputs and the researchers and organizations involved, thereby promoting collaborative and reproducible science. Its comprehensive features support the entire research lifecycle, from data organization and annotation to sharing and publication, aligning with the principles of open and FAIR (Findable, Accessible, Interoperable, Reusable) science.

SEEK employs the ISA (Investigation, Study, Assay) infrastructure, a standardized framework that organizes individual experiments into broader studies and investigations. This structure is adaptable beyond biological sciences, accommodating various research domains. The platform's **data catalogue** includes raw datasets, Standard Operating Procedures (SOPs), models, publications, and presentations, all organized by projects and linked to their respective contributors. This **organization** enhances data discoverability and contextual understanding.

Recognizing the dynamic nature of research data, SEEK offers **flexible sharing permissions** that align with the research lifecycle—from initial collection and collaborative phases to public dissemination upon publication. Researchers can control **access** at granular levels, ensuring appropriate data sharing at each stage. The platform supports versioning and assigns static URLs to data items, facilitating proper citation and credit. Upon publication, SEEK can generate Digital Object Identifiers (DOIs) for individual items or aggregated research objects, enhancing the traceability and impact of research outputs.

SEEK incorporates semantic technology to enable sophisticated content queries. Through integration with RightField, a tool for embedding ontology **annotations** in Excel spreadsheets, researchers can semantically enrich their data, improving interoperability and reuse. The platform allows in-browser **exploration** and annotation of Excel spreadsheets without requiring downloads, streamlining data examination and collaboration.

For computational models adhering to the Systems Biology Markup Language (SBML) standard, SEEK provides **in-platform simulation** capabilities, facilitating immediate model validation and analysis. Additionally, the platform supports **interlinking of data, models, and publications**, offering a comprehensive view of research projects and their interconnected components. This feature aids in understanding the relationships between different research outputs and the overall investigative context.

SEEK maintains an index of researchers, projects, and organizations, fostering a **collaborative environment** where users can identify potential collaborators and experts. Profiles detailing areas of expertise enable efficient networking and knowledge exchange, strengthening the **research community**.

The platform is available for **local deployment**, allowing research groups to host their own instances tailored to specific needs. Alternatively, researchers can register on the FAIRDOMHub, a **publicly accessible instance** of SEEK maintained by FAIRDOM for community use. This flexibility ensures that SEEK can accommodate diverse research environments and collaboration scales.

19 Simulation platform – the architecture

19.1 High-level architecture

Building on the developments and considerations outlined in the previous chapters, a high-level overview of the Platform architecture is proposed in the figure below. The platform is implemented as a collection of microservices to deliver a Platform as a Service (PaaS) solution for managing Digital Models and executing simulation jobs. The goal is to evolve the platform to provide Trusted Research/Execution Environment (TRE/TEE) services, fully support Digital Twins and implement federated Simulation/ML algorithms.

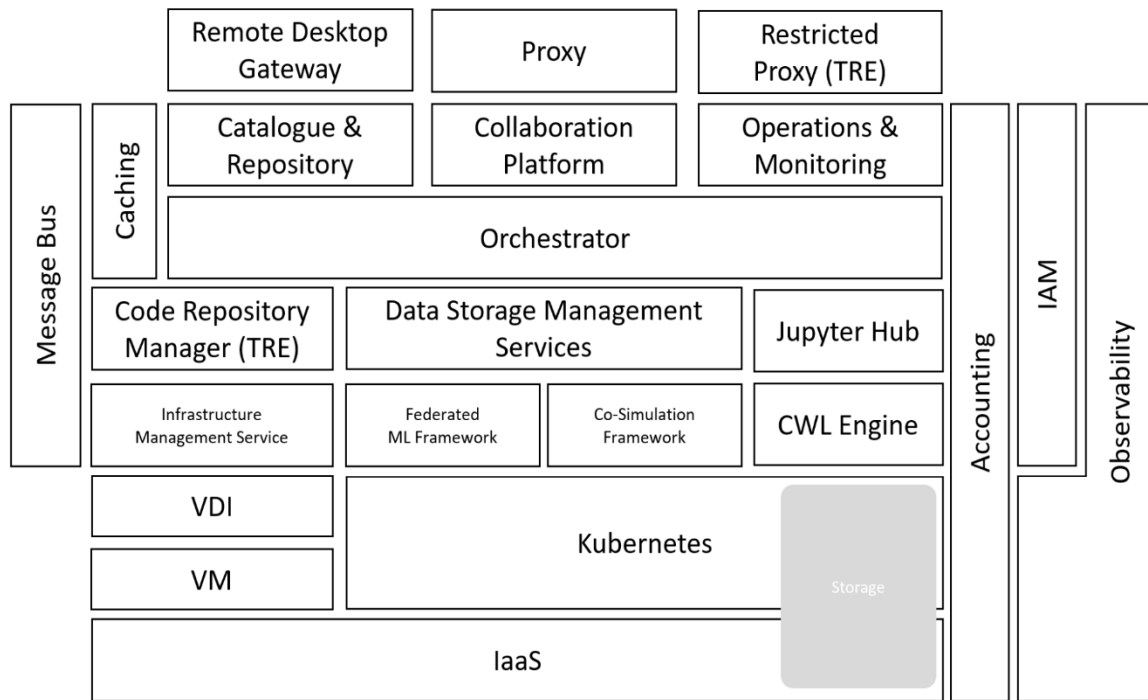


Figure 22: High-level overview of the proposed VHT Platform Architecture.

The main components of the proposed architecture are:

1. **IaaS:** the Infrastructure as a Service (IaaS) component is responsible for managing physical resources such as compute, storage and network. A custom solution like OpenStack can be deployed on premises or a 3rd party cloud solution may be used such as Microsoft Azure, Amazon AWS or Google GCP. The IaaS will provide customized Virtual Machines (VM), Virtual Desktop Infrastructure (VDI) for running desktop workloads and applications and a Kubernetes cluster for deploying our PaaS solution.
2. **Kubernetes:** on top of the IaaS component, a Kubernetes cluster can be installed on which the PaaS solution will be deployed. All software components will be executed as containers in this cluster.
3. **Storage:** the deployed microservices may require disparate storage types such as a block or an object storage system. Storage will be configured both at the level of IaaS and Kubernetes based on the operational requirements of each service. For instance, a relational database may require local storage provisioned directly from the IaaS layer.
4. **IAM:** the Identity & Access Management component is responsible for user authentication and authorization and enforcing security policies across all services. Application specific roles required by the Platform's services will be configured in the IAM component.
5. **Observability:** this component consists of all services and agents responsible for monitoring the execution and performance of all PaaS components and Infrastructure. It will collect logs files and metrics, store it and provide the tools to analyse it. The data collection will be performed depending on each component's implementation, using either external agents or through programmatic integration. Both pull and push models will be considered for the data collection.

6. **Accounting:** the accounting component is responsible for monitoring the cost of use of the Platform's resources and the execution of models (credits) and the remuneration to model, data and infrastructure providers (billing). It will be able to compute charges based on historic data from the observability component as well as to make estimates based on compute, storage and network requirements. In the first phases, the Platform could support the use of credits to model the remuneration for using the Platform's resources. Afterwards, an accounting component can be integrated with a billing component to offer commercial services. Furthermore, both components could be integrated with the billing APIs of existing cloud providers to offer suggestions about the cost of executing a service on a 3rd party cloud solution.
7. **Message Bus:** since the implementation of the VHT's PaaS solution is based on a microservice architecture, a message bus will be used to provide reliable and asynchronous communication to the services, assuring a low coupling between the deployed services.
8. **Caching:** to improve the overall system performance, the caching component will allow any service to cache application specific data. Caching will save both compute and network resources and will improve the end-user experience.
9. **Remote Desktop Gateway:** the platform is going to provide VDI solutions, *i.e.*, customized virtual machines for desktop workloads and applications. To avoid exposing specific protocols and services like Remote Desktop (RDP), VNC or X11 forwarding over SSH, the remote desktop gateway can provide access to a VDI instance using any modern web browser with support for HTML5 features.
10. **Proxy:** the proxy component is the entry point to the PaaS applications and services.
11. **Restricted Proxy:** this component is similar to the proxy component but it is instead used by the Trusted Research Environment components to assure that only network traffic to safe servers is allowed. For instance, the restricted proxy will only allow pulling Docker images from specific organizations' image registries or checking source code from specific code repositories.
12. **Catalogue & Repository:** this component provides services for storing, discovering and sharing digital models, datasets, notebooks and any artifacts related to digital models and simulation. Data published to the Catalogue has unique identifiers and are discoverable by third party services using established mechanisms such as the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).
13. **Collaboration Platform:** the collaboration platform offers the end-users office and document collaboration tools, Enterprise File Synchronization and Sharing (EFSS) services and communication applications. Availability of the EFSS services alone, accessible through the user dashboard module of the Operations & Monitoring component, can be a starting point. Later, additional tools can be added such as generation of automated metadata for selected resource types.
14. **Operations & Monitoring:** the catalogue and collaboration platform are operating mostly independently. Operations and Monitoring (OM) provides all the value-added services the platform offers such as CWL workflow execution or infrastructure resource provisioning. Whenever an end-user makes a request, the OM component will propagate it to the appropriate components as required. It will also collect data from several systems to monitor the progress of the request execution. The OM component consists of the following three modules:
 - **User Dashboard:** the user dashboard is the main entry point to all Platforms services for the authenticated users. It provides the User Interface (UI) to access the EFSS services, configure and submit jobs and visualize their results.
 - **Admin Dashboard:** This module is used by administrators to monitor the status of the deployed microservices and to perform common administrative tasks such as user and role management. The admin dashboard provides only a higher level overview of the Platform's status, if more detailed information is required, the observability component should be used instead.
 - **REST API:** The Representational State Transfer (REST) API is used to implement the Platform's federation capabilities. It allows the communication among internal components, such as the orchestrator and data storage management services, of remote clusters.
15. **Infrastructure Management Services:** the Infrastructure Management Services (IMS) component will allow users to request infrastructure resources or applications from the IaaS or Kubernetes

components *e.g.*, a VM. The IMS will also provide an abstraction layer over commercial cloud solutions such as Microsoft Azure or Amazon AWS.

16. **Data Storage Management Services:** the platform will consist of several services, each one with its one storage requirements. For instance, the catalogue service may store datasets as S3 objects while the collaboration platform may store user files using NFS. The Data Storage Management Services (DSMS) will decouple applications that require access to files from different services by replicating required files as needed and also forcing any applicable security policy.
17. **CWL Engine:** the CWL Engine component is responsible for executing CWL workflow instances. The orchestrator component will manage the data input file replication by interacting with the data storage management services, it will apply any security policies and will monitor the workflow execution.
18. **Jupyter Hub:** this component is used to support Jupyter Notebooks by the platform .
19. **Co-Simulation & Federated ML Frameworks:** the VHT Platform should be integrated co-simulation and federated ML frameworks. This component will handle file access, security enforcement and invocation of any selected framework.
20. **Orchestrator:** whenever a user requests a task, *e.g.*, the provisioning of a VM or the execution of a CWL workflow, a job is scheduled at the orchestrator service. The latter will instrument all the required services like the DSMS and CWL Engine to execute and monitor all scheduled jobs reporting progress to the end-users through the Operations & Monitoring component.

19.2 Components and technologies used in EDITH Proof-of-Concept (PoC) Infrastructure

During the EDITH-CSA project, a Proof-of-Concept Infrastructure, the EDITH Infrastructure, was created to test a number of the aforementioned architectural choices and inform the final architecture proposed in this Roadmap.

Box 32: EDITH development – Proof of Concept Infrastructure

EDITH development: VHT Proof-of-Concept Infrastructure

Status: MVP#1 deployed on Athena resources: <https://dashboard.edith.athenarc.gr/home>

The box below highlights the components that have been implemented in the first minimum viable product release (MVP#1). For components that have been instantiated using existing third party implementations, the figure shows the logo of a service candidate that may be used for the actual deployment.

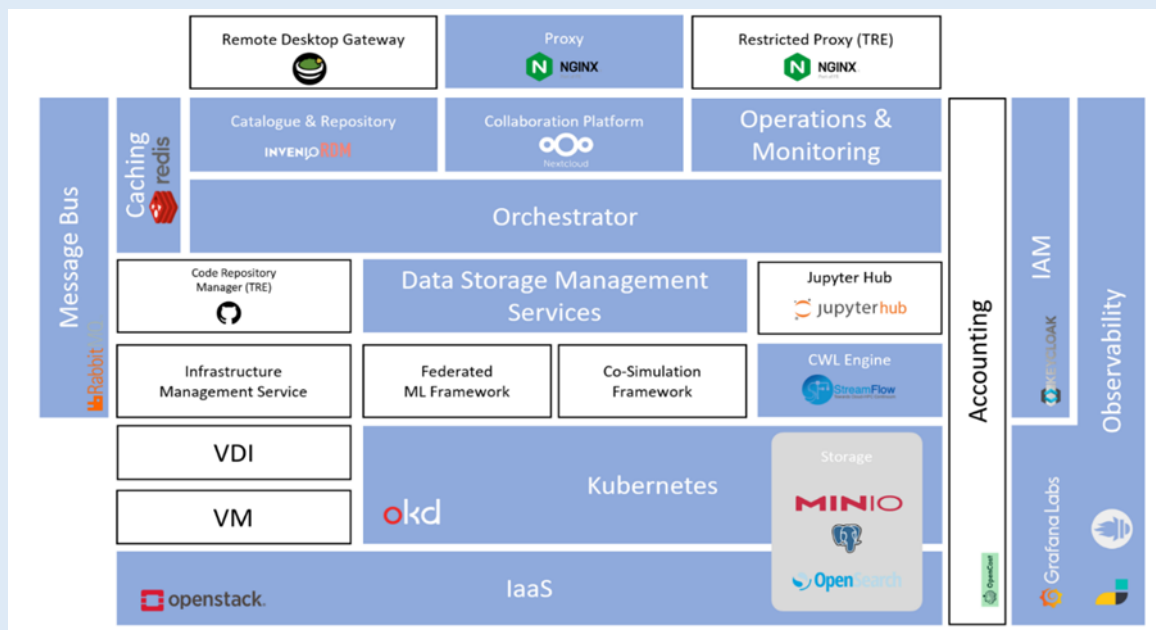
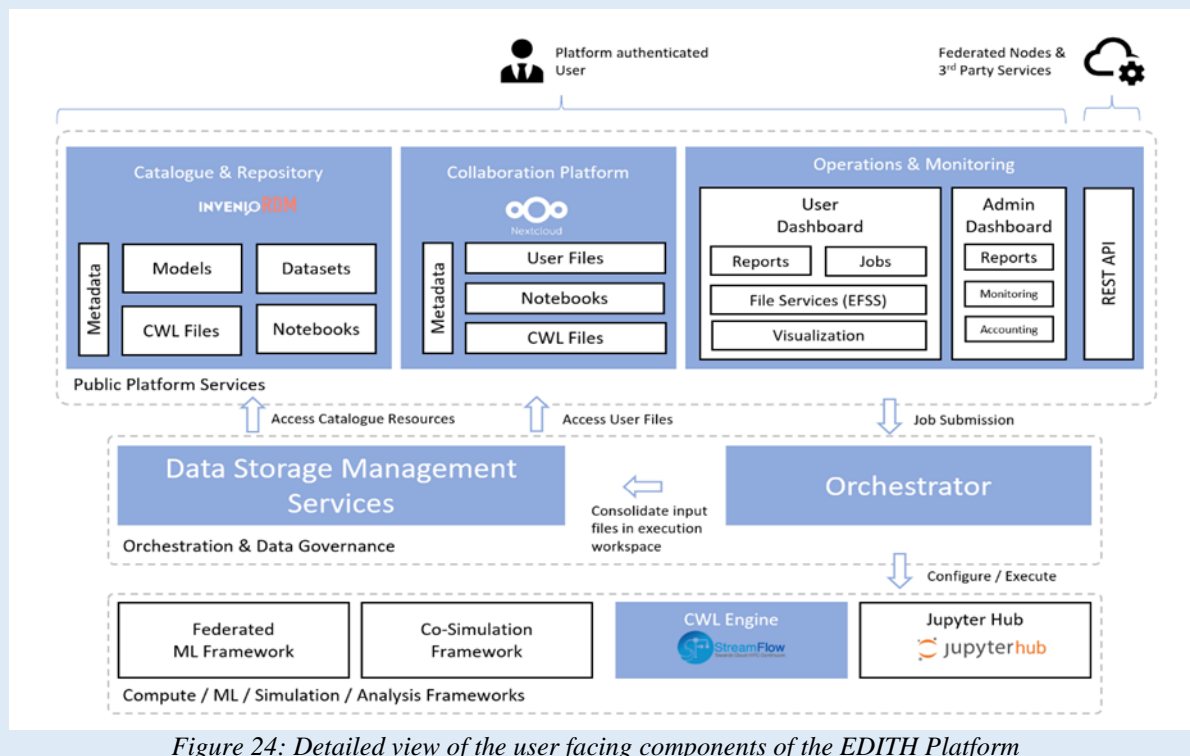


Figure 23: Components of MVP#1

Below is a more detailed view of the user facing components of the Platform, *i.e.*, the components that directly interact with users, including the two main components that interact with them, namely the Data Storage Management Services and the Orchestrator. For simplicity, the proxy services, the components responsible for VM provisioning and the infrastructure components are not displayed.



19.3 Computational architecture

As previously mentioned, the Infrastructure as a Service (IaaS) component is responsible for managing physical resources such as compute, storage and network. The minimally required computational architecture consists of the following elements and layers.

19.3.1 Hardware - physical layer

The physical layer of the computational architecture consists of the (virtual) machines that are available to run computational tasks.

High-end supercomputing clusters are integral components of this layer, offering a large number of nodes and CPU cores. HPC infrastructures are usually operated by dedicated computing centres, many of which work on a non-profit basis and receive public funding in order to support (among others) scientific research projects. Thus, they are often well integrated with the broader European data and computing infrastructures. Accelerator devices (GPUs) may be available on all or a subset of nodes. Notably, there exists a very fast interconnect between nodes, cores, and GPUs. Additionally, **interactive access**, visualization, and login nodes contribute to the user interface. High-performance internet access is facilitated from login and visualization nodes, although direct internet access from compute nodes is typically restricted. Furthermore, these clusters feature very large and **fast file systems**. Operating as multi-user systems, they are organized using Unix user IDs and groups. Direct access is achieved through SSH, often incorporating two-factor authentication. Programmatic access is also enabled via REST APIs (*e.g.*, UNICORE). The MPI is supported, allowing for the execution of very large-scale jobs. Accounting mechanisms are in place, utilizing projects and grants for tracking usage.

Cloud resources, another facet of this layer, consist of small-to-medium-sized server systems, often comprising a single node with multiple cores. GPUs may be available within these cloud environments. Typically, these resources are partitioned into single-user virtual machines using Platform as a Service (PaaS) software like OpenStack or Kubernetes. Cloud resources boast good to excellent internet

connectivity. They are particularly useful in supporting the so-called “long-tail science” where “classical” HPC resources are not regarded as necessary, as is the case for many types of VHT simulations and data processing tasks.

A separate category within this layer includes **devices, instruments, and edge (fog) computing resources**. Primarily designed for data collection, these resources do not foresee computational workloads and are not explored further in this context. Development of suitable access interfaces for end users (many of whom are not IT experts) is an important aspect of this resource. Edge/Fog computing is the concept of having a processing node which is topologically proximate to end-user devices, which is useful in situations which call for low latency or even real-time processing of data, such as, for example, data that is generated by sensors. In such circumstances, for applications deployed in a medical care setting, it might be useful to deploy a local compute node, capable of processing urgent requests, while more computationally heavy tasks are delegated to external HPC or Cloud resources. Such local nodes are often operated - on a limited scale - by medical care providers, and can be leveraged in the context of VHT applications for persons receiving treatment at a given location.

All systems capable of executing computational tasks are collectively referred to as “**execution sites**”. Properly describing computational sites involves metadata, such as size (number of nodes, of total cores, of GPUs), type (multi-node batch, single-node, *etc.*), inter-node connectivity, internet connectivity, and file system capabilities (size, throughput, quotas, *etc.*). This metadata is crucial for informed decision-making regarding the allocation and execution of tasks on specific execution sites.

19.3.2 Software layer

The software layer, a critical component of the computational infrastructure, is comprised of both the environment (including operating systems, system software, administrative tools, *etc.*) and the specific applications utilized for running computational tasks.

Recognizing the anticipated heterogeneity of the infrastructure, the recommendation is to achieve fully **containerized application software**. Containers, lightweight virtualized software environments, enable the packaging and execution of software in a manner that is nearly independent of the host system. This level of abstraction enhances portability, requiring minimal special software installation on the host system. A container runtime, such as Docker or Apptainer, is the primary requisite.

In exceptional cases, driven by performance considerations on HPC systems, it may be necessary to run “native” software – traditional installations via software modules. In such instances, efforts should be directed towards packaging or wrapping the software to conceal system peculiarities as much as possible.

Descriptions of applications, including CWL workflows, necessitate **metadata for informed decision-making regarding the required resources**. This includes:

- Required system size (number of nodes, total cores, GPUs, *etc.*);
- Required system type (multi-node batch, single-node, *etc.*);
- Required inter-node connectivity;
- Required internet connectivity;
- Required file system capabilities (available size, throughput, *etc.*).

Machine-readable descriptions of a computational task will be a (set of) CWL documents and related files. To facilitate execution, such descriptions should be bundled into an archive file (*e.g.*, zip) to be sent to the execution framework. Importantly, task descriptions must remain independent of the computational site. Additionally, data sets, ideally described using a system-independent approach such as a DOI or other persistent identifier, must be accessible to the application during runtime.

Computational tasks are **executed through a container runtime**, like Docker or Apptainer. Each instance of a computational task will be run in its own working directory (“sandbox”) to prevent unintended side effects, such as interactions with concurrently running tasks. The runtime mounts the working directory and any required data directories into the execution environment. To mitigate side

effects, directories are generally mounted in a read-only mode, except in cases where write access is absolutely necessary (*e.g.*, the working directory itself).

19.3.3 Access layer / Access APIs

The access layer is that part where the generic computational task is sent to a particular execution site, the correct user ID, group and accounting project is selected, and the task is initiated.

HPC access, typically facilitated through SSH terminal access, poses challenges when considering the use of "service accounts" for storing and deploying required credentials. This approach often proves impractical and may conflict with HPC centre security policies and usage agreements. A more viable solution involves the use of a fully-fledged federation software, such as UNICORE²⁸⁶. UNICORE offers advanced features for job submission and management, as well as integration with authentication and authorisation infrastructures. Widely available at European HPC centres, notably at Jülich, UNICORE effectively addresses access issues. When alternative access tools may exist at different sites, the software layer needs to adapt to and utilize them.

At a minimum, the following functionalities are essential:

- Automatic user ID mapping based on the provided means of authentication;
- Capability to launch tasks based on the abstract task description;
- Monitoring of tasks;
- Integrated tools for data access, upload, and download in a consistent manner with computational tools.

On HPC systems (cloud platforms), computational **resources** are allocated to users based on **project grants**. For each execution site, the system (or cloud platform) must be aware of available grants and allow users to make selections. Users may possess private grants, and the system should support the utilization of these as well. The system (or cloud platform) must also handle **accounting**, ensuring that a single user cannot exploit the system by consuming more resources than appropriate. Robust measures are essential to prevent unauthorized consumption and maintain fair resource distribution in alignment with project grants and usage policies.

Box 33: EDITH development – Model Execution Environment

EDITH development : VHT demonstrators in Model Execution Environment

Status: *deployed on ACC Cyfronet HPC resources*

To turn the idea of the Virtual Human Twin into practice, an appropriate software environment is required, backed up by modern compute and storage resources. We have analysed the internal structure and functional requirements of typical applications simulating human physiology, among others in the context of osteoporosis, cardiovascular, cancer and glyceimic control. This formed the basis for elaboration of a demonstrator of the execution subsystem of the VHT ecosystem. The demonstrator is a software platform which facilitates repeatability, replicability, and reproducibility of computations. It is based on the Model Execution Environment (MEE) developed in EurValve, PRIMAGE, and InSilicoWorld EU projects.

The demonstrator is agnostic to the supported classes and formats of data items. It supports a comprehensive data repository where various data items may be queried, retrieved, and fed into the computational models which constitute the simulation workflow, to run on classical HPC resources for scale-out studies which involve processing large amounts of data and "parameter study" types of computations. It enables model versioning (past versions of the model are stored and may be referred to if needed) as well as reproducibility of computer simulations. It also permits execution of computational models via a set of scripts, with a versioning system which enables collaborative editing and tagging specific versions that may be later selected to suit the researchers' needs. It provides a straightforward way to display, download and analyse simulation results. Short-lived ssh certificates are used for user rights delegation.

The demonstrator hides the complexity of the underlying computing infrastructure and introduces a unified way for data to be maintained. The functionality of the demonstrator was successfully validated with a set of available typical VHT modules on ACC Cyfronet HPC resources. The demonstrator's feature set is extendable.

²⁸⁶ <https://www.unicore.eu>

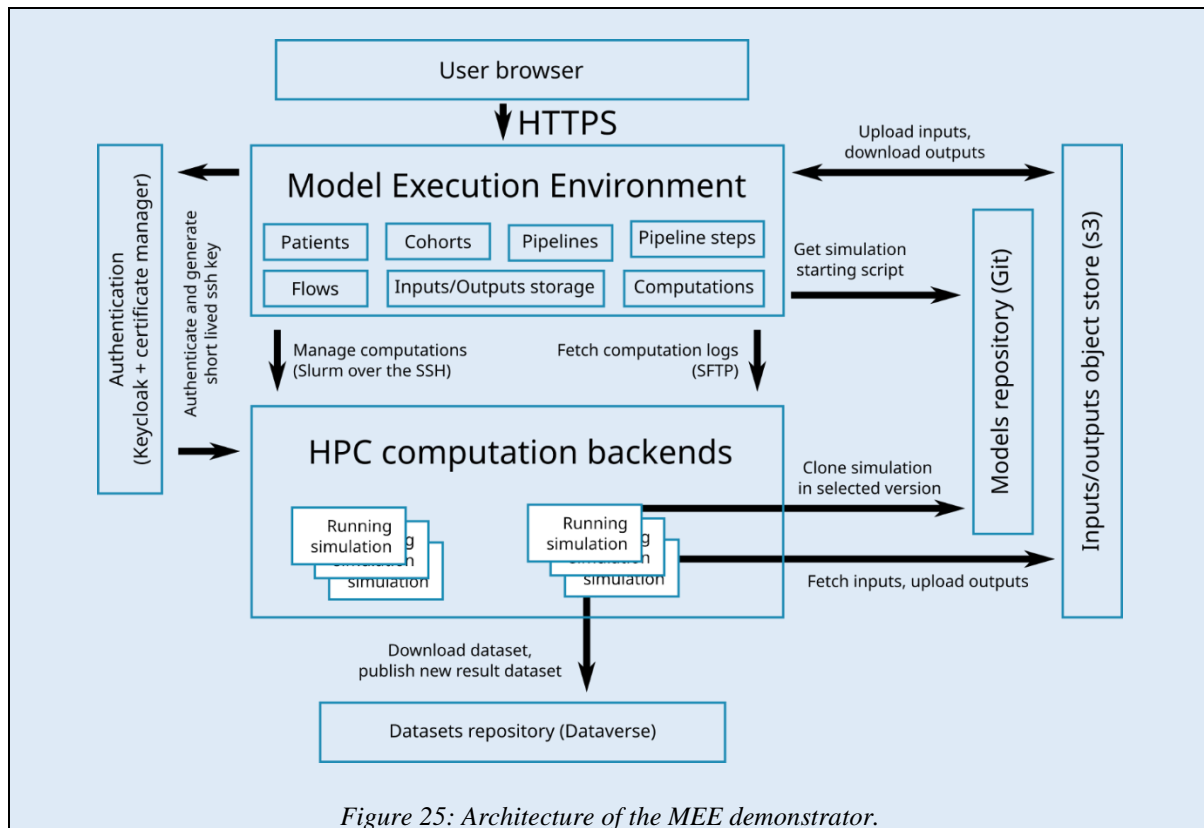


Figure 25: Architecture of the MEE demonstrator.

19.4 Testing and Quality Assurance

19.4.1 Testing

Since the VHT platform is to be expandable by design, ensuring interoperability between different software and services, and integrating with various tools and systems, adequate testing processes are required. Where needed, specific guidelines are to be provided by the VHT to potential contributors. A concise summary of testing types is discussed here, to give an overview of the scope and scale of potential testing strategies.

Unit-tests refer to testing each individual module and unit of the source code of a component in order to determine that they perform as expected. A unit may be an individual function, method, procedure, module or object. Unit testing helps identify and fix errors at an early stage, ensuring that each unit functions as expected. In addition, it is necessary to run **integration tests** which verify the overall functionality of the system, by testing each new module which is to be integrated into the EDITH RI. This step ensures that different modules function together properly. It detects issues arising from the integration of components, such as data flow problems or communication errors.

End-to-end (e2e) testing can be divided into two categories: system tests and acceptance tests. The **system** testing assesses the entire system's compliance with specified requirements, identifies issues that may arise from the interaction of different subsystems and ensures the system behaves as expected. The **acceptance** testing validates whether the system meets the end-users' requirements, involving tests performed by end-users or stakeholders to verify that the system is ready for deployment.

Performance testing (load testing, stress testing, and scalability testing) assess the system's responsiveness, scalability, and stability under various conditions. It ensures the system can handle expected workloads. **Security testing** (vulnerability scanning, security audits) enhances the system's resilience to cyber threats and safeguards sensitive data, by identifying vulnerabilities and weaknesses in the system's security. Finally, with **usability testing**, we evaluate the system's user-friendliness and user experience. We want the system to be intuitive and easy-to-use.

19.4.2 *Quality Assurance*

Dedicated **quantitative and qualitative metrics** will need to be established, along with quality assurance procedures, to ensure that the VHT platform meets defined standards, aligns with requirements, and delivers reliable and high-quality outcomes. Extracted requirements, encompassing both functional and non-functional aspects, will be utilized to comprehend **user expectations**, and metrics will be formulated, covering response times, system availability, and **user satisfaction**. The testing strategies mentioned above will be employed to enhance efficiency.

Monitoring is required to track the platform's performance, promptly identify potential issues, and ensure operation within set parameters. Regular checks will be conducted to verify compliance with relevant regulations, standards, and legal requirements.

An integral aspect of quality assurance is **user involvement**. Comprehensive **documentation**, including user manuals and technical guides, will be provided to offer clear guidelines for platform usage and maintenance. **Training sessions** need to be made available for both existing and new users, regardless of their expected role in the platform.

Establishing feedback loops with users, developers, and other stakeholders is crucial. **Regular feedback collection** will inform ongoing improvements and address emerging requirements, primarily facilitated through the introduction of new use cases.

20 Digital Twin use cases

Box 34: Use Case - Glycemic control in ICU patients (workflow).

Use Case: Glycemic control in ICU patients

Status: start-up company, solution in clinical trial

Website: www.insilicare.com

InSiliCare's AI-powered Digital Twin is a clinical decision support solution for ICU patients requiring glycemic control. It provides personalized, safe, and effective management of blood glucose levels and nutrition delivery. Insulin and nutrition treatments are calculated to maximise safety from hypoglycemia, while controlling patient blood glucose levels and optimizing nutrition towards a configurable physician-determined practice of care.

The model integrates patient data and model parameters to dynamically update insulin sensitivity profiles in real-time. Hosted on a secure virtual machine within the hospital infrastructure, it ensures all calculations remain local, safeguarding patient data. The software supports EHR integration for automated data input but also allows manual entry. Upon receiving new data, the Digital Twin updates automatically, generating tailored treatment recommendations and setting a follow-up measurement schedule once validated by a clinician. Notifications ensure clinical staff are prompted for timely glucose measurements.

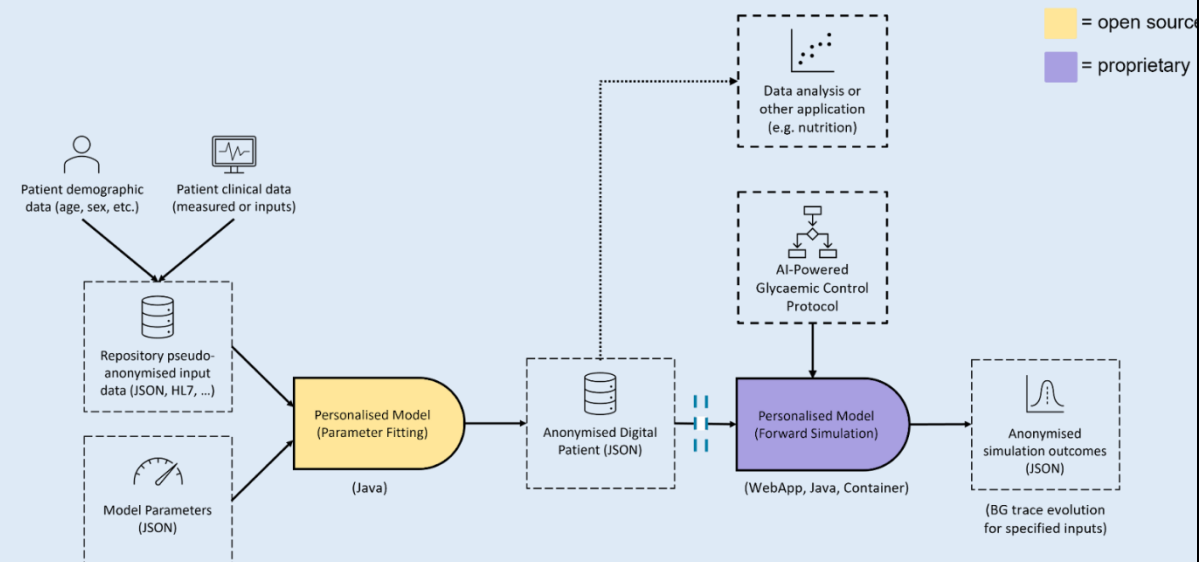


Figure 26: This figure depicts data workflow for the complete GC solution to be commercialized by Insilicare. In the context of EDITH, the workflow stops at the Anonymised Digital Patient step (dashed blue lines).

Box 35: Use Case - Osteoporotic fracture risk prediction (workflow).

Use Case: Bologna Biomechanical Computed Tomography for osteoporotic fracture risk prediction

Status: clinical research studies

Bologna Biomechanical Computed Tomography (BBCT) is a Digital Twin methodology designed to predict the mechanical strength of the femur under critical loading conditions in osteoporotic patients. Quantitative Computed Tomography (QCT) scans of the hip region and patient data inform a subject-specific Finite Element (FE) model able to predict the risk of hip fracture at the time the CT is performed (ARF0).

The computational model is developed through several steps (Figure 27), performed both on the operator's local workstation (e.g., Step 1) and on an HPC cluster (e.g., Steps 2–4):

1. Image processing and anatomical landmarking. CT scan images are semi-automatically segmented to extract the geometry of the femur. The operator semi-automatically identifies specific anatomical landmarks, which are used to define local reference systems and loading conditions.
2. Model Generation and stochastic fall generation. The main steps required to construct the finite element (FE) model are performed. These include mesh generation, calibration, and mapping of material properties. One million potential fall scenarios are then simulated based on a body with the

height and weight of the patient. The resulting impact forces are estimated. Variables—such as initial and final velocities, accelerations and impact attenuation—are modelled as normal distributions informed by literature data.

3. **FE Simulations.** The patient-specific model is run 28 times, varying the femur's orientation at the impact. To replicate a sideways fall, a concentrated force is applied at the centre of the femoral head, while a contact interaction is defined between the surface of the greater trochanter and a rigid plane oriented perpendicular to the force. The distal part of the femur is constrained to a hinge located at the knee's centre.
4. **Result Analysis.** For each impact orientation, the force intensity required to fracture the femur is computed based on principal strains in a region of interest located proximally. By comparing the FE-derived failure loads to the estimated impact forces, the absolute risk of hip fracture (ARF0) is calculated.

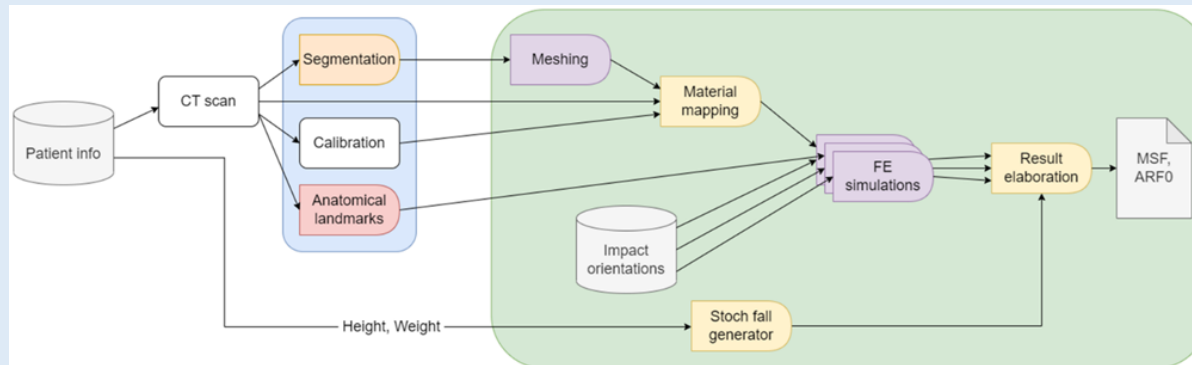


Figure 27: High-level workflow of the BBCT solution

Box 36: Use Case - Universal Immune System Simulator for Tuberculosis (workflow).

Use Case: Universal Immune System Simulator for Tuberculosis

Status: *augmenting clinical trials*

Website: <https://www.mimesis.srl/>

The Universal Immune System Simulator (UISS) offers a robust, clinically validated platform to address these challenges by simulating and predicting immune responses *in silico*. UISS-TB has demonstrated its utility in optimizing therapeutic strategies for Tuberculosis, evaluating drug regimens, and predicting outcomes.

The UISS workflow comprised of the following steps:

1. **Data Integration:** Patient-specific data, including cytokine levels and bacterial load, are fed into the model.
2. **Simulation:** UISS-TB dynamically updates immune response profiles based on input parameters.
3. **Output:** Tailored treatment recommendations are generated, including follow-up schedules and immune response predictions.
4. **Validation:** Outputs are reviewed by clinicians to guide decision-making and optimize therapy.

Box 37: Use Case - Epileptogenic zone localisation for surgical planning in epilepsy patients (workflow).

Use Case: Epileptogenic zone localisation for surgical planning in epilepsy patients

Status: *in clinical practice*

Website: <https://www.cloudsofcare.com/>

Persyst ESI powered by Epilog is a neuroimaging solution that automatically combines scalp EEG data with a patient's MRI to perform Electrical Source Imaging (ESI). It pinpoints the origin of brain activity linked to seizures, helping clinicians accurately localize the epileptogenic zone—critical for surgical planning in epilepsy patients.

The workflow consisted of the following steps:

1. **Data Collection:** EEG recordings capture electrical activity from the brain using standard electrode arrays.
2. **Integration:** The EEG data is paired with high-resolution MRI scans to map electrical signals onto the patient's brain anatomy.

3. **Analysis:** Advanced algorithms calculate the probable source of the electrical activity within the brain, visualizing it as a 3D map.
4. **Interpretation:** This map aids clinicians in identifying the specific brain region responsible for seizures, even in cases where traditional imaging (*e.g.*, MRI alone) is inconclusive.

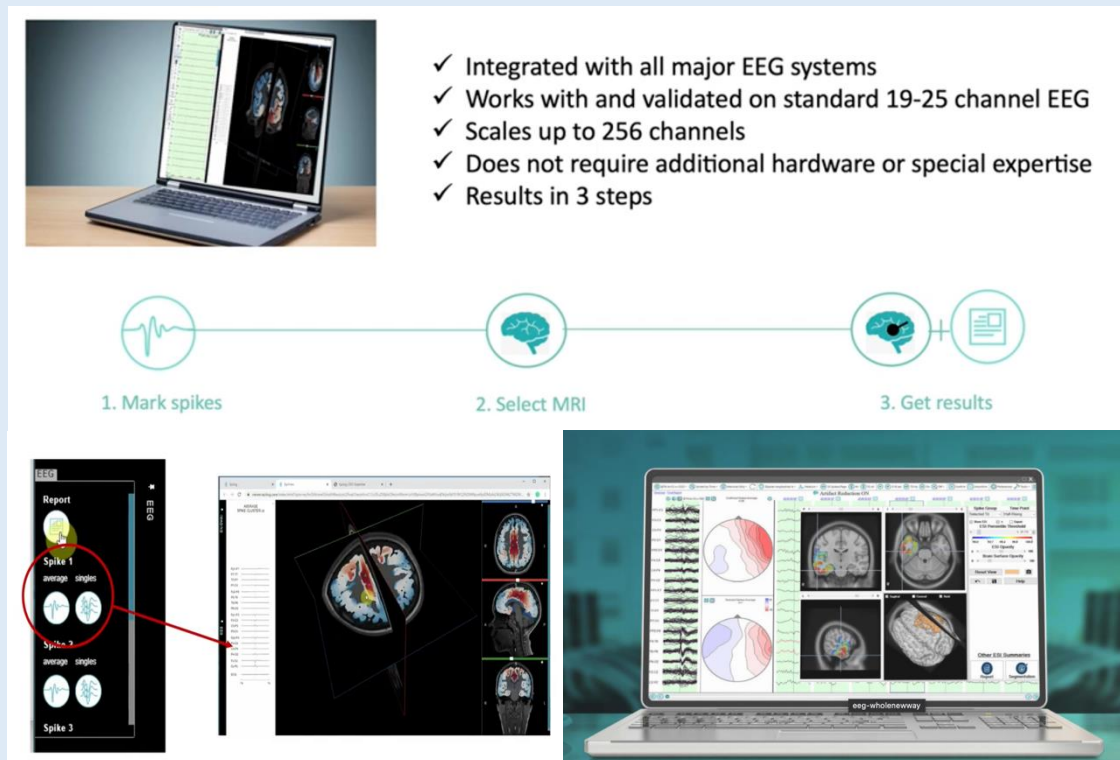


Figure 28: Top: overall workflow for Persyst ESI powered by Epilog. Bottom: details of the clinician's interface.

Box 38: Use Case - the Atrial Modelling Toolkit for cardiovascular Digital Twins (workflow).

Use case: the Atrial Modelling Toolkit for cardiovascular Digital Twins

Status: for research purposes

Website: <https://github.com/pcmlab/atrialmtk>

The Atrial Modelling Toolkit²⁸⁷ (**atrialmtk**) aims to overcome the challenges of constructing cardiac models at scale through the development of a robust, open-source pipeline for bilayer and volumetric meshes for atrial models.

Different workflows can be built with the atrialmtk toolkit, two such workflows are discussed here.

Workflow 1 takes a segmentation mask of the atrial blood pool as input to define the endocardial surface and generate a bilayer atrial mesh for simulation. The segmentation can be produced from the raw image data (*e.g.*, CT, MRI) by an expert, or input directly by the user if already provided, including through automated segmentation.

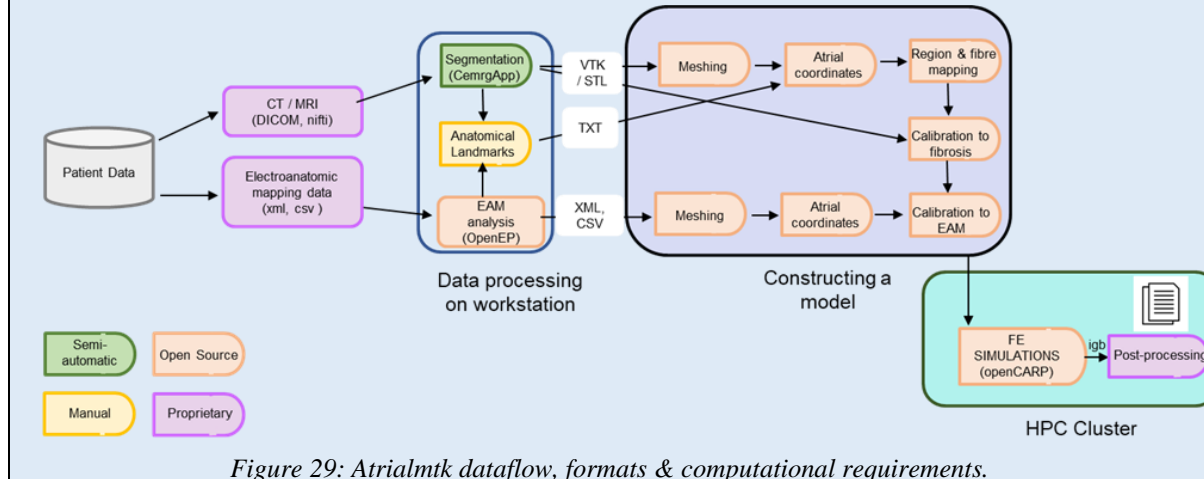
Workflow 2 allows the user to process an existing atrial mesh, that has both surface and volumetric data, but from which the pulmonary veins, left atrial appendage, and vena cava have already been clipped. This is often the case when using artificial datasets, for example, those produced by statistical shape models. Since it starts from atrial mesh data in the absence of any atrial vessels, it does not include the ImagingData or Clipping steps from workflow 1, and can be used to generate bilayer or volumetric atrial meshes.

General aspects of the use case's workflows and dataflows.

- **Data needs:** To construct an anatomical atrial model, imaging (MRI or CT) or electroanatomic mapping data is required.

²⁸⁷ <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2023.0038>

- **HPC needs:** To run cardiac electrophysiology simulations (*e.g.*, OpenCarp²⁸⁸), HPC resources are required. A typical simulation would be about 1 hour on 64 cores.
- **Data outputs:** Meshes will be stored in .vtk format, with associated vector fields for atrial fibre directions, and scalar fields for any extra measurables (for example, tissue conductivity values, late-gadolinium enhancement MRI intensities). Simulation experiment files for the opencarp solver are also available²⁸⁹.



²⁸⁸ <https://opencarp.org>

²⁸⁹ <https://www.sciencedirect.com/science/article/pii/S0169260721002972>

21 Infrastructure for VHT: Conclusions and Recommendations

21.1 Conclusions

Part 3 of the roadmap dives into the crucial infrastructural elements required to support the development, deployment, and sustainable operation of the VHT. This part emphasizes the need for a robust and secure infrastructure capable of handling the complexities of data storage, model integration, and collaborative research within the VHT ecosystem.

To realize the ambitious goals of the VHT, a sophisticated infrastructure comprising three interconnected software components is required: the **Catalogue, the Repository, and the Platform**. The Catalogue serves as the central hub for discovering and accessing VHT resources, providing detailed metadata and facilitating efficient resource retrieval. Complementing the Catalogue, the Repository provides secure and reliable storage for the diverse range of VHT resources, including data, models, workflows, and annotations. The Repository ensures data integrity, version control, and facilitates seamless data sharing among authorized users. Finally, the Platform acts as the interactive workspace where users can analyse, simulate, visualize, and interact with VHT resources. The platform provides a suite of tools and services, enabling users to perform complex computations, execute simulations, and collaborate on VHT-related research.

Central to the VHT infrastructure is the establishment of a **robust and comprehensive Catalogue and Repository** system. This system will not only enable efficient resource management but also ensure FAIR (Findable, Accessible, Interoperable, and Reusable) principles are adhered to. Achieving this requires adopting a standardized approach to metadata annotation, utilizing community-agreed ontologies and controlled vocabularies. The Catalogue and Repository system must also incorporate robust security measures to protect sensitive data, ensuring compliance with data privacy regulations and fostering trust among users.

A well-defined **governance framework** is crucial for overseeing the decision-making processes related to populating the VHT with data and model objects. This framework will establish clear guidelines for data submission, metadata annotation, and credibility assessment, ensuring the quality and trustworthiness of the VHT resources. A transparent and standardized procedure for adding new resources to the VHT is essential for fostering community engagement and promoting the adoption of the VHT.

The realization of a fully functional VHT simulation platform requires careful consideration of both its **architectural design** and its **integration with existing digital/compute infrastructures**. The platform should leverage federated technologies and adopt a modular design to ensure scalability, flexibility, and efficient resource allocation. The platform must also provide robust security measures, ensuring data privacy, integrity, and compliance with regulatory requirements. Integrating the VHT platform with existing high-performance computing infrastructures will be crucial for enabling complex simulations and handling the computational demands of large-scale VHT applications.

21.2 Recommendations

Developing a **robust VHT infrastructure requires a multi-faceted approach**, encompassing technical expertise, community engagement, and policy considerations. Addressing the challenges of data standardization, interoperability, security, and governance will be crucial for realizing the transformative potential of the VHT and enabling its widespread adoption in healthcare. As such, following recommendations can be made:

1. **Designing the VHT resource Repository and Simulation Platform:** The design, construction, and continuous enhancement of the VHT resource repository and simulation platform are crucial aspects. It is imperative that this platform adheres to all applicable laws and regulations in Europe, particularly concerning data privacy and security. The platform should leverage distributed and federated computing, incorporating regional or country-specific nodes to ensure data sovereignty

and efficient resource utilization. A federated authentication and authorization mechanism is essential to manage access securely and facilitate collaborative data sharing across the VHT ecosystem. In addition to security, the platform should prioritize interoperability, modularity, flexibility, scalability and extensibility to accommodate future growth and advancements in VHT technologies. Embracing the principles of Open Science, Open Source, and Open Data will foster collaboration, accelerate innovation, and enhance transparency.

2. **Investing in a robust and adaptable IT infrastructure:** To ensure the longevity and continuous growth of the VHT, sustained support for the development, testing, and implementation of advanced and interoperable IT platform architectures is needed. This includes investments in computational infrastructure, cybersecurity, HPC, cloud services, and edge infrastructure to store, manage and process the vast amounts of data generated by the VHT. Continuous inclusion and updates of domain-specific services are crucial to enable advanced workflows and applications, provide tools for seamless integration, harness new technologies, and visualise results.
3. **The future of the VHT - data availability and access:** For the VHT to reach its full potential, access to high-quality, annotated, and interoperable digital health data is paramount. The VHT initiative should work in conjunction with existing and developing infrastructures and platforms, like the European Health Data Space. Standardized formats and semantics will be essential for interoperability and effective data analysis. Patient privacy, personal data, health, and safety must be safeguarded, which can be achieved through anonymization, pseudonymization, and synthetic data generation. Data curation, certification, and quality control will ensure the reliability of the data, with data stewards and curators playing key roles.
4. **Co-evolution of the VHT - a collaborative ecosystem for success:** The VHT should not exist in isolation but actively co-evolve with existing digital services and research infrastructures. Embracing a global vision for the European life sciences research infrastructures, the VHT initiative should establish strong links to existing platforms, including ESFRI and EOSC, to avoid duplication and foster synergy. The VHT should leverage existing efforts and available platforms, mapping existing modelling and computational infrastructures, both commercial and open source, to ensure efficient resource utilization and maximize impact. Long-term support is crucial, requiring engagement from national networks and continuous collaboration with other initiatives to create a thriving ecosystem for the advancement of the VHT.
5. **Fostering collaboration - co-designing the VHT for optimal usability:** To ensure the successful adoption of the VHT technology, a co-design approach including end-users and other stakeholders is essential. Interactive design sessions can foster inclusion, allowing stakeholders to provide feedback and contribute to the development of a user-friendly platform. This user-centric design places ease of use at the forefront through documentation, demos, and training materials tailored to specific user needs. A strong emphasis on user experience and usability testing will ensure the platform's functionality aligns with the needs and expectations of diverse stakeholders. The development of an automated one-stop-shop catalogue can streamline access to resources and best practices, reducing barriers between research and practice. This participatory approach emphasizes communication and transparency by using clear language to communicate design processes, composition, and best practices to user groups.

PART 4:

**REALISING THE VHT – ELSI, STANDARDS &
REGULATORY**

22 Standards for VHT

As repeatedly discussed in all previous PARTS of this roadmap, Adopting or establishing common standards and protocols for data sharing and model development ensures seamless data integration across national borders and facilitates the creation of comprehensive VHT models.

The standards are defined by **Standard Defining Organizations** (SDOs) such as ISO (*e.g.*, ISO/TC 215 “Health Informatics”, ISO/TC 276 “Biotechnology” and ISO/TC 194 “Biological and clinical evaluation of medical devices”), CEN/CENELEC, IEC (*e.g.*, IEC/TC 62 “Medical equipment, software, and systems”) and national counterparts. They play a crucial role in collecting the expertise from all different domains relevant for the VHT and in maintaining the standards on a long run, but **scientific community standardization initiatives** like COMBINE or GA4GH are needed as well to drive the standardization of novel VHT and modelling technologies and approaches (see Table A.12 and A.13 in Annex 2).

What follows in this chapter is an **overview of the currently available standards**, as well as relevant ones that are currently in development. The final choice of the optimal standards to be used will depend on the choices made in the tech stack during the implementation phase of the VHT. A more comprehensive and detailed listing of the recommended standards for the VHT can be found in Annex 2 of this roadmap and as a searchable online listing in the EDITH FAIRSharing collection²⁹⁰. More detailed descriptions of the standards are given in the “Standardization landscape, needs and gaps for the virtual human twin (VHT)”²⁹¹ and “EDITH standards implementation guide (IG)”²⁹² documents²⁹³. The chapter

Standards for data formats, data integration and data input into models

Standardisation of modelling

Standards for metadata of data and models – semantic annotation and taxonomy

Semantic annotation & taxonomy for vht

22.1 Standards for data formats, data integration and data input into models

22.1.1 Standards for data formats, integration, interoperability and access

The ISO 20691:2022²⁹⁴ standard “*Requirements for data formatting and description in the life sciences*” requires that the data in the life sciences shall be FAIR (Findable, Accessible, Interoperable and Reusable). Therefore, the data used in models and simulations of the human Digital Twin should at least encompass the information described in the minimum information standards and follow the FAIR²⁹⁵ principles. For being FAIR, the data must have a unique ID, must be linkable, must have an assigned licence and must be annotated by metadata describing for instance the disease, the tissue, the cell type and the used modelling parameters.

Other requirements described in ISO 20691:2022 are about the consistent formatting and documentation of data, models and metadata as well as the requirements concerning storing, sharing, accessing, interoperability and reuse of data, models and metadata in the life sciences.

The ISO 20691:2022 standard acts as a **reference framework or hub standard** for other life science data and integration standards. It describes the requirements and rules for applying standards for formatting, description and documentation of data types in the life sciences and contains a catalogue of criteria and requirements for interoperable life science data formats and semantic data description standards.

²⁹⁰ <https://fairsharing.org/4787>

²⁹¹ <https://zenodo.org/records/10492796>

²⁹² <https://zenodo.org/records/10524795>

²⁹³ <https://fairsharing.org/4787>

²⁹⁴

<https://www.iso.org/standard/68848.html#:~:text=This%20document%20specifies%20requirements%20for,human%20biological%20research%20and%20development>

²⁹⁵ <https://www.go-fair.org/fair-principles/>

The annex A of ISO 20691:2022 contains a list of **common recommended formats** for life science data and the annex B contains a list of **minimal reporting standards** for data, models, and metadata. This set of metadata describes the context of the data and models.

The ISO 20691 **FAIRSharing collection**²⁹⁶ contains a listing of these standard formats, which should be used to represent life science data and to annotate them by ontology terms. This list encompasses data formats for -omics, biochemical and molecular biology methods as well as formats for biological imaging and for computer models of biological systems. That FAIRSharing collection is actively managed and maintained, so that it is always on the last stand.

A further standard is ISO/TS 9491²⁹⁷ “*Biotechnology — Recommendations and requirements for predictive computational models in personalised medicine research*”, which consists of two parts:

- Part 1: Guidelines for constructing, verifying, and validating models;
- Part 2: Guidelines for implementing computational models in clinical integrated decision support systems.

Finally, ISO standard ISO 4454:2022²⁹⁸ “Genomics informatics – Phenopackets: A format for phenotypic data exchange” standardised the description of phenotypic information and the ISO 23494 series²⁹⁹ “Biotechnology – Provenance information model for biological material and data” describes how to document the provenance information for biological data and samples.

There are also some **missing standards**, *e.g.* there is no standard for model benchmarking. For that it would be desirable to have model testing sets as gold standards for the different modelling and simulation tools, as well as testing tools, like for instance the SBML test suite³⁰⁰. Another missing standard is one for granting and controlling access to data for model validation and instantiation (with patient data). Beside the standard for model quality, a standard for data quality (for construction, simulation and validation data) and for its assessment is missing, even if this aspect is partially addressed by ISO/TS 9491-1 and 9491-2). Typical model and data quality criteria are listed also in Table 11. One big problem is how to operationalize these quality criteria.

Data integration combines data from different sources like Electronic Health Records (EHRs), life-style data, environmental and registry data, molecular -omics data, laboratory data, imaging data, biosignal, intensive care vital signs data and other medical relevant sorts of data. Data integration can be done either on the individual level or on the variable level. Individual level integration is required for personalised models and means that all available data for an individual patient are brought together for the model building and simulation tasks.

It was demonstrated that it is possible to integrate several modelling frameworks (agent-based models, ordinary differential equations, stochastic reaction systems, constraint-based models, solid-body physics and spatial diffusion) into a composite model. But a standard for such an integration of different models into a multiscale model is still missing, even though there are some proposals for such integrative frameworks^{301,302}.

The ISO/14199³⁰³ standard “*Health informatics — Information models — Biomedical Research Integrated Domain Group (BRIDG) Model*” is a domain model for **data interchange** that enables **semantic interoperability** and intends to bridge between biomedical, clinical research and routine healthcare data. It can be seen as a meta-standard and can be used as a starting point and blueprint for the integration of modelling and simulation data with other health data and how to make them interoperable.

For the interoperable exchange of health record data the Fast Healthcare Interoperability Resources (FHIR) from HL7 is the most common standard, and observational data are preferentially encoded in OMOP-CDM.

For a comprehensive list of further clinical interoperability standards see Table 12.

²⁹⁶ <https://fairsharing.org/3533>

²⁹⁷ <https://www.iso.org/standard/83516.html>

²⁹⁸ <https://www.iso.org/standard/79991.html>

²⁹⁹ <https://www.iso.org/standard/80715.html>

³⁰⁰ <https://sbml.org/software/sbml-test-suite/>

³⁰¹ Agmon *et al.*, Bioinformatics, 2022 ;38(7),1972–1979, <https://doi.org/10.1093/bioinformatics/btac049>

³⁰² Masison *et al.*, PNAS. 2021; 118 (20) e2024287118. <https://doi.org/10.1073/pnas.2024287118>

³⁰³ <https://www.iso.org/standard/83433.html>

Currently there are **no harmonised strategies for data access**, since the data protection laws are varying between the countries and can even vary from federal state to federal state, as in Germany. What shall be used are strategies to control the access to restricted (*e.g.*, person-related) data by researchers. Such harmonised Data Access Agreements (hDAAs) are still missing now.

22.1.2 Standards for human data collection and formatting

For **medical imaging** there are three important standards. The first one is the Digital Imaging and Communications in Medicine (DICOM), which is the most used data structure in health care, and which stores the data as 2D layers. The Brain Imaging Data Structure (BIDS) is a group of standards defined by the International Neuroinformatics Coordinating Facility (INCF) and is mainly used in neuroscience/neuroimaging research. For several special imaging techniques extensions to BIDS are available, which contain mainly metadata describing features for these special techniques. A third group of standards were defined by the Neuroimaging Informatics Technology Initiative (NIFTI) and stores the data in a true 3D volume format. Table A.14 in Annex 2 gives an overview about these formats and their extensions/descendants.

Since the annexes of ISO 20691 and ISO/TS 9491 list mainly systems biology formats, in the following some **formats for physiology/biomedicine** are mentioned. There is a plenitude of formats in the area of **electrophysiology, biosignal and time series of vital sign data** (*e.g.*, body temperature, systolic and diastolic blood pressure, pulse rate, respiration rate, oxygen saturation, BMI...) recorded in intensive care. Many of them are proprietary formats defined by medical device manufacturers. Therefore, in Table A.15 of Annex 2 only formats supported by Standard Defining Organisations (SDOs), community standards, or formats, which are at least de facto standards, are listed.

For encoding of **genomic sequence variants**, one should prefer a data format mentioned in the ISO 20691 document, if applicable. An overview about sequence variation formats is given in Table A.16 in Annex 2.

In molecular tumour boards and computational oncology, such formats for sequence alignment, genome and copy number variation and DNA methylation patterns are used. These are often stored in the NCI Genomic Data Commons (GDC) or the Portable Format for Bioinformatics (PFB) format. PFB is based on the Apache Avro serialisation format and able to encapsulate the data model, the data dictionary, the data itself and controlled vocabularies in one file³⁰⁴.

BioCompute objects (BCO) is the IEEE 2791-2020 standard for documenting workflow steps of bioinformatics analyses of next-generation sequencing (NGS) and high-throughput screening (HTS) and is used for regulatory submissions to the FDA.

22.1.3 Standards for clinical practice

To optimise the research and patient treatment, some **minimal clinical datasets** specific for a disease were defined. For example, for cancer there is the mCode³⁰⁵ (minimal Common Oncology Data Elements) HL7-STU (standard in trial use) available. It describes the minimum information, which should be collected in the electronic health records of cancer patients. Also, for some other diseases, *e.g.*, for multiple sclerosis³⁰⁶, Parkinson disease and stroke, the definition of minimal clinical datasets is available (see common data elements catalogue³⁰⁷).

For some diseases XML-based standard file formats based on such disease-specific common data models for capturing the relevant information are available. Examples are cMDX (Clinical Map Document based on XML) for prostate cancer or BRCAPRO for breast cancer.

There are a lot of reporting, best practices/**clinical practice guidelines** (CPGs) and consensus statements. Most of them are specific to the different medical fields. These best practice guidelines are a foundational pillar for evidence-based medicine and guide clinicians in their daily work decisions.

³⁰⁴ Lukowski *et al.*, PLoS Comput Biol. 2023; 19(3):e1010944. <https://www.doi.org/10.1371/journal.pcbi.1010944>-

³⁰⁵ <http://hl7.org/fhir/us/mcode/>

³⁰⁶ <https://medical-data-models.org/19067?form-lang=en>

³⁰⁷ <https://www.commondataelements.ninds.nih.gov/cde-catalog>

Many of these guidelines are related to the use of artificial intelligence (AI) in medical diagnosis, treatment, or prognosis. Others like PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) are reporting guidelines for systematic reviews of a field. Despite these guidelines are very domain specific, there are some general statements and rules defining what should be reported in these guidelines, *e.g.* by the PICAR (Population, Intervention, Comparison, Attributes of eligible CPGs, Recommendation characteristics) and the PICOR (Population, Intervention, Comparator, Outcome and Recommendation) statements.

These clinical practice guidelines are also used for meta-analysis. The quality of observational studies is directly influenced by the quality of the captured data. Therefore the capturing of the meta-variables defined in these guidelines should be done in a standardized and curated way, so that high quality is ensured. Otherwise the findings of the studies can be easily biased. Often not only the value of the variable itself, but also the severity and duration of how long the value was present, or even the fluctuation of the variable value over time, are important and should be captured in a standardized way.

International Patient Summary (IPS) is a set of basic patient-related physiological and clinical data and the ISO 27269:2021, EN 27269:2022 and CEN/TS 17288:2020 standard. It comprises data about medications, allergies/intolerances, problems, immunizations, results and procedures for a specified patient. It is a joint standard of five standard defining organisations (CEN, HL7, IHE, ISO and SNOMED) and actively supported by the Global Digital Health Partnership (GDHP) and the World Health Organisation (WHO). It can be implemented either by HL7 V3 Clinical Document Architecture (CDA), with HL7 Fast Healthcare Interoperability Resources (FHIR) or according to the Integrating the Healthcare Enterprise (IHE) IPS profiles.

Phenopackets³⁰⁸ version 2 is a Global Alliance for Genomics and Health (GA4GH) and ISO 4454³⁰⁹ **open standard for sharing disease and phenotype information** of a patient/sample, especially in the context of diseases, especially for rare diseases, cancer, but also for common diseases. A Phenopacket links detailed phenotypic descriptions with disease, patient, and genetic information, diagnosis, and treatments. The phenotypic features are *e.g.*, signs, symptoms, laboratory and imaging findings, behavioural manifestations, or the results of physiological tests.

The Phenopackets schema is specified as a *protobuf* schema³¹⁰. These protocol buffers are platform- and language-neutral, so that serialisation code for several programming languages can be generated easily. Therefore, Phenopackets is a standard for the language- and platform-neutral exchange of phenotypic information.

22.1.4 Other data format types

If one has specific modelling requirements, then other data format types than the ones listed above are needed. Examples are FASTQ, a sequencing format with quality information. SAM (Sequence Alignment Map), BAM (Binary Alignment Map) and CRAM (Compresses Reference-oriented Alignment Map) are file formats for alignments.

For microscopy data OMEX (Open Microscopy Environment XML) is a standard for storing acquisition parameters, annotations, and image analysis results of microscopy experiments and BDML (Biological Dynamics Markup Language) is a standard format for storing spatiotemporal dynamics of molecules and cells in live imaging data.

For data from wearable health devices and trackers, *e.g.*, weighing scales, blood pressure monitors, biosensors like blood glucose monitors, smart health watches, and mobile ECG devices for arrhythmia recognition, the ISO/IEEE 11073 standard for personal health devices (PHD) shall be followed.

22.1.5 Standards for the description models

For **AI-based models** there are no standardised file formats for the input data available. Since these models are based on big data, they often use hierarchical (.hdf, NetCDF), row-based (.csv, .json), columnar (*e.g.*, Apache Parquet) or other high-performance (*e.g.*, Apache Iceberg) formats as input format for the data used for AI-/ML-based learning.

³⁰⁸ <http://phenopackets.org>

³⁰⁹ <https://www.iso.org/standard/79991.html>

³¹⁰ <https://developers.google.com/protocol-buffers>

Most **ML frameworks** define their own model exchange format. For a standardised and interoperable representation and deployment of ML and deep learning (DL) models, one of the Open Neural Network Exchange (ONNX), Portable Format for Analytics (PFA) or Predictive Model Markup Language (PMML) standards can be used. Since PFA is more flexible than PMML and easier to extend to new algorithms and model types and also integrates necessary pre- and post-processing steps, PFA or ONNX shall be used.

The structure and parameters of trained neural networks can be stored and exchanged with the Neural Network Exchange Format (NNEF) defined by the Khronos group.

IEC SC42 defines standards for artificial intelligence and the use of synthetic data for artificial intelligence methods³¹¹.

According to the Report “*Artificial Intelligence Standardisation Landscape Update*”³¹² the following 3 standards are very relevant with respect to the European AI Act:

- IEEE P7003: Standard for Algorithmic Bias Considerations;
- IEEE P7001/D4: Draft Standard for Transparency of Autonomous Systems;
- IEEE 7000: Standard Model Process for Addressing Ethical Concerns during System Design.

A detailed table of AI standards can also be found in the Report “*Landscape of AI Standards*”³¹³ of the Technical Working Group AI of the CSA project StandICT.eu

For **mesh-based finite element modelling (FEM)**, boundary element method (BEM) or finite volume modelling (FVM) often the .vtk (Visualisation ToolKit) output file format is used. For the FEM input files FieldML may be used. For the generation of 3D-models from DICOM data, *e.g.*, for computer guided implant surgery, standard tessellation language (STL) files shall be used, which allow the printing of 3D models.

Models, which are not stored in a standard format and encoded in a scripting language, *e.g.*, in Matlab/Octave, Python, R, Julia scripts, ..., can be executed on the platform either if they have an **interface** that allows the import/export, respective conversion into one of the aforementioned standard formats, or if the **scripts are directly executed inside the workflow**. Otherwise, if the scripts are “injected” into the workflow, they can be generically packed into a Binary FHIR resource.

For some kinds of models, *e.g.*, cellular automata there are no explicit standards available, but at least some of these models can be described by using the MultiCellML and Morpheus standards. For others, like Petri net models there is the ISO/IEC 15909 standard, but this standard is not biology specific.

For versioning of models the PAV ontology, which is a specialization of W3C-PROV-O, shall be used. A standard for the life-cycle management of the DTs in the VHT, their components, models and the data belonging to a patient is still missing, even if for some aspects one can build upon the patient-centred clinical decision support framework proposed by Sittig et al.³¹⁴. Clinical risk management is regulated by the ISO 14971:2019 standard (“Medical devices - Application of risk management to medical devices”).

22.2 Standards for the modelling process

The **modelling process itself** shall follow (where possible) the ISO/TS 9491-1³¹⁵ technical standard “*Biotechnology - Recommendations and requirements for predictive computational models in personalised medicine research - Part 1: Guidelines for constructing, verifying and validating models*”. In order to standardise the modelling process, all data, metadata, models and simulation results shall be documented according to the FAIR³¹⁶ and ALCOA³¹⁷ (Attributable, Legible, Contemporaneous, Original, Accurate) principles. Therefore, one shall use standard formats with semantic annotations based on defined ontologies.

³¹¹ <https://iec.ch/blog/isoiec-report-address-synthetic-data-techniques-used-chatgpt-and-other-ai-tools>

³¹² <https://publications.jrc.ec.europa.eu/repository/handle/JRC131155>

³¹³ <https://zenodo.org/record/5011179>

³¹⁴ <https://pubmed.ncbi.nlm.nih.gov/37414544/>

³¹⁵ <https://www.iso.org/standard/83516.html>

³¹⁶ <https://www.go-fair.org/fair-principles/>

³¹⁷ <https://www.eurotherm.com/life-sciences-cpg/data-integrity-life-sciences/alcoa/>

ISO/TS 9491-1 specifies requirements and recommendations for **models used for research purposes in the field of personalised medicine**, *i.e.* computational models used in routine clinical, diagnostic or therapeutic purposes are excluded. In detail ISO/TS 9491-1 contains specifications for:

- the design, development, and establishment of predictive computational models.
- the set-up, formatting, validation, simulation, storing and sharing of computational models.
- data used to construct or be required for validating such models.
- formatting, descriptions, annotations, interoperability, integration, access and provenance of such data.

To standardise the **modelling process**, the data going into a model shall be integrated and the modelling results shall be validated. The integration of data means the systematic combination of different data necessary for the modelling task and belonging to a patient.

The **validation** means the comparison between the output of the calibrated model and the measured data.

The annex A of ISO/TS 9491 contains information on common standards relevant for personalised medicine and *in silico* approaches. Part 2 of ISO/TS 9491 contains recommendations for implementing computational models into clinical decision support systems.

22.2.1 Data preparation

The modelling process and the different types of models are described in detail in the ISO/TS 9491-1³¹⁸ document “*Guidelines for constructing, verifying and validating models*”, that describes in detail the steps of **data preparation for integrating the data into the models**.

- Sampling the data;
- Data formatting and harmonisation, *e.g.*, lab values have a unit associated with them. This unit can either be mass/volume or mol/volume. Therefore, the values shall be converted to a unique scale. For that the molecular weight of the analyte must be known;
- Data description by descriptive metadata, describing for example the context of the datasets;
- Semantic annotation of the data, *e.g.*, by annotating genes and proteins with ontology terms;
- Definition of a data interoperability framework;
- Data integration, either on the personal or on the variable level;
- Adding data provenance information;
- Defining who can access the data.

22.2.2 Standards for Models

As discussed in previous PARTS, one can distinguish **knowledge-driven models** based on prior knowledge and hypotheses on the causal relationships (mechanistic models), and **data-driven models** such as Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) models.

Typical examples for models are:

- the risk prediction for common diseases;
- disease course and therapy response prediction;
- pharmacokinetic/-dynamic modelling and *in silico* trial simulations;
- data-driven artificial intelligence (AI) models.

For computer models of biological/biomedical systems **a selection of the following standards** should be used. They are often defined by community efforts like the COMBINE³¹⁹ consortium and listed in detail in the Annex A of the ISO/TS 9491-1 document “*general recommendations and requirements for modelling in personalised medicine*” as well as in the Annex A and Annex B of the ISO 20691 document “*Biotechnology — Requirements for data formatting and description in the life sciences*”. They are also registered in different registries like for instance the BioSimulators FAIRsharing collection³²⁰, the Bio-simulation website³²¹ and the Normsys registry³²².

The most important standards in this category are listed below:

³¹⁸ <https://www.iso.org/standard/83516.html>

³¹⁹ <https://co.mbine.org>

³²⁰ <https://fairsharing.org/1536>

³²¹ <https://docs.biosimulations.org/>

³²² <https://normsys.h-its.org>

- Systems Biology Markup Language (SBML), an XML-based description and exchange format for differential-equation models of biological processes. Level 3 of SBML is a modular format that allows to define extension packages for other than differential-equation models;
- CellML, an XML-based description and exchange format for cellular models;
- Biological Pathways eXchange (BioPAX) for the exchange and visualisation of biological pathway data.
- Synthetic Biology Open Language (SBOL) describing the exchange of synthetic biological genetic parts, devices, modules, and systems;
- Neuroscience eXtensible Markup Language (NeuroML) is an XML-based description and exchange format for models in neuroscience consisting of the 4 parts Biophysics, ChannelML, MorphML, and NetworkML;
- The Pharmacometrics Markup Language (PharmML) is an exchange format for pharmacokinetic and pharmacodynamic models.

A more extensive list of formats for biomedical computer models is given in Table A.17 in Annex 2. For multicellular models there are currently 3 standards (MorpheusML, MultiCellDS and MultiCellML plus some proprietary formats, *e.g.* from CompuCell3D). A COMBINE network group was established, which works on uniting these standards into one standard for virtual tissue models (OpenVT), so that a repository of multi-cellular models can be established in a similar way as SBML has spurred the BioModels repository for differential equation-based models.

Standards for model simulations and documentation of results

For the simulation of systems biology models one of the standard formats listed in Table A.18 in Annex 2 shall be used. Here the most important formats are:

- SED-ML (Simulation Experiment Description) for describing the setup of simulation runs;
- COMBINE Archive or OMEX (Open Modelling Exchange) for storing all simulation-related files in a compressed format;
- SBMRL for encoding SBML simulation results.

Standards for Graphical Model Visualisation

- KEGG (Kyoto Encyclopaedia of Genes and Genomes) Markup Language (KGML);
- Systems Biology Graphical Notation (SBGN), a graphical notation for representing biological processes;
- Systems Biology Open Language (SBOL-Visual) is a graphical notation to specify genetic parts, devices, modules, and systems.

22.3 Standards for metadata of data and models – semantic annotation and taxonomy

It has been estimated that 30% of the world generated data comes from the healthcare sector. Despite the considerable potential offered by all this data, **proper access and re-use remains a huge challenge**, for many reasons. The main reason seems to be the lack of a standardised approach to represent such information. This poses several challenges.

- Datasets coming from different sources cannot be combined;
- The meaning of variables contained in the dataset might be not understandable;
- Unstructured data cannot be efficiently used to feed models, as, beside the variable per se, other pieces of information might be lacking, *e.g.*, unit of measurements;
- Variables within a dataset are usually not systematically related one to each other, losing in this way valuable information along the way.

The **use of a common basis for data representation** is essential to meet the FAIR principles, in particular the concept of interoperability³²³. Within the VHT, there is an exquisite variety in technologies to generate data but also create and solve models, that is representative of the heterogeneity of the questions that need to be answered in human health. This further exacerbates the already challenging situation mentioned above and several sub-communities have started trying to organise themselves (*e.g.*, logical modelling³²⁴). Furthermore, all the above points apply not only to human intelligence, but also to “machine intelligence”³²⁵. In fact, applying a standard method for representing data would facilitate the development of automated data processing. One solution is represented by implementing a semantic annotation procedure fully tailored to the needs of the VHT.

³²³ Guizzardi G, Data Intell.(2020);2(1–2): 181–191. https://www.doi.org/10.1162/dint_a_00040.

³²⁴ Touré V *et al.*, Briefings in Bioinfo (2021); 22(4): bbaa390. <https://doi.org/10.1093/bib/bbaa390>

³²⁵ Schneider T. and Šimkus M. KI – Künstliche Intelligenz (2020); 34(3):329–353. <https://www.doi.org/10.1007/s13218-020-00686-3>.

Semantic annotation is the process to ascribe a specific meaning to elements in a dataset or following a structured and standardised manner, usually based on ontologies, being collections of terms, relational expressions, and associated natural-language definitions designed to capture the intended interpretation of these definitions (Figure 8). Ontologies are necessary when it comes to multi-scale representations, since they can provide a taxonomy and, thus, specifying how the variables are related one to each other. There are two ISO documents^{326,327}, which recommend the **use of ontologies and standard terminologies** for the descriptions of entities and concepts. They also provide recommendations on how to maintain an ontology (*i.e.*, versions control) and the minimum information to report.

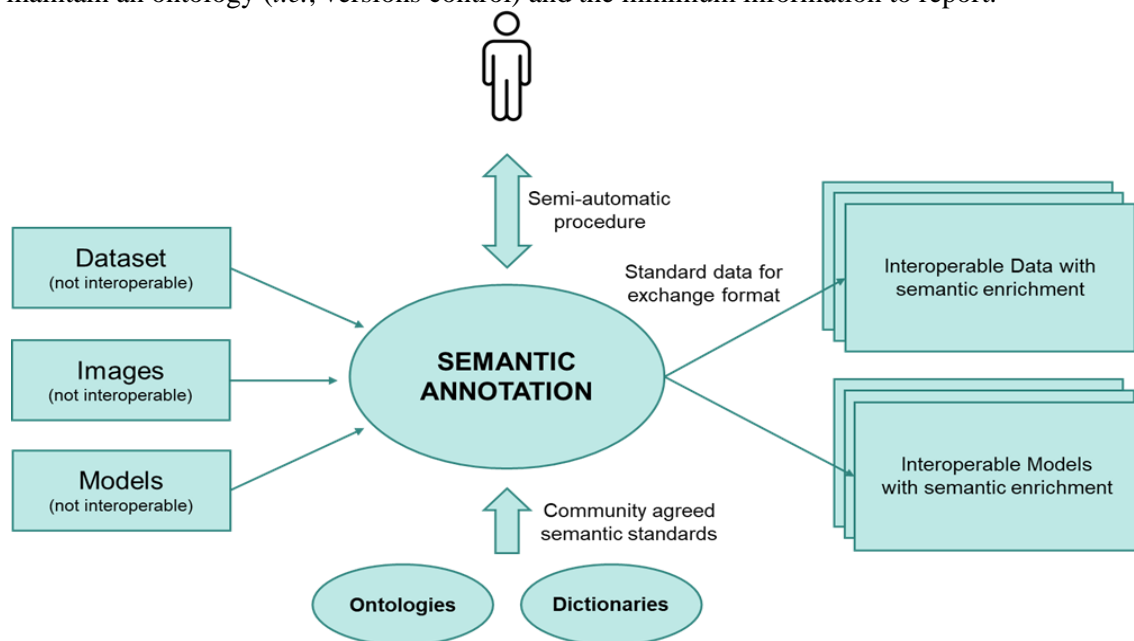


Figure 30: Semantic annotation. Adapted from Sasse J et al.³²⁸.

22.3.1 Metadata requirements

Annotation of DTs and their parts, components and entities with metadata is a prerequisite to make the models and their entities FAIR. For integrating data from heterogeneous sources the often different naming and coding conventions of metadata attributes must be taken into account. Examples of heterogeneous sources with different conventions are systems like FHIR (Fast Healthcare Interoperability Resources), openEHR (open Electronic Health Records), OMOP (Observational Medical Outcome Partnership) and the CDISC (Clinical Data Interchange Standards Consortium) standards SDTM (Study Data Tabulation Model) and CDM (Common Data Model). For mapping all metadata attributes from all source formats, an intermediate convergence format (“metadata crosswalk”), which includes the maximum set of metadata attributes, can be used³²⁹. Another challenge is that data fields are often populated with values of different data types. Sometimes data fields from the input format must be split or merged to achieve a mapping to the target format.

The ISO 20691:2022 standard “*Requirements for data formatting and description in the life sciences*” provides **recommendations and requirements for the semantic description and annotation of data** in the life sciences (including construction and validation data for computational models and DTs). This includes a set of minimum consensus information for the annotation of biological data. Metadata and the annotation of the data integrated into DHTs should also consider domain-specific minimum information guidelines. A complete list of minimum information guidelines is given in Annex B.2 of

³²⁶ ISO 20691:2022 for “Requirements for data formatting and description in the life sciences”

³²⁷ ISO/TS 9491-1 for “Recommendations and requirements for predictive computational models in personalised medicine research”

³²⁸ Sasse J et al. Appl. Sci 2022; 12(2):796. <https://www.doi.org/10.3390/app12020796>

³²⁹ Bönnisch et al., Sci Data 2022;9, 659. <https://doi.org/10.1038/s41597-022-01792-7>

ISO 20691:2022 “*Biotechnology — Requirements for data formatting and description in the life sciences*”.

Specifically for predictive computational models in personalised medicine the ISO standard ISO/TS 9491-1 “*Biotechnology — Recommendations and requirements for predictive computational models in personalised medicine research — Part 1: Guidelines for constructing, verifying and validating models*” comprises recommendations and requirements for the semantic description and annotation of the models and data that is integrated into them.

For systems biology models, MIRIAM (Minimum Information Requested in the Annotation of biochemical Models) and for simulations MIASE (Minimum Information About a Simulation Experiment) could be relevant. MINIMAR (Minimum Information for Medical AI Reporting) is a proposed standard for reporting of healthcare prediction models based on artificial intelligence methods. A minimum information guideline for annotating **parameter identifiability and model inference** – something like a Minimal Information for Model Inference and Parametrization (MIMIP) - is still missing³³⁰. PETab is a format for specifying parameter estimation problems in systems biology in a tabular format.

The metadata attributes and entities should be described by using **domain-specific standard terminologies**, controlled vocabularies and ontologies as defined in the Annex B.3 of ISO 20691:2022. Such terms describe for instance the species, sex, age, organ, tissue, cell type, disease, and so on. Annotations for identifiable objects or processes, manipulated entities or used technologies are also possible. Typical metadata for modelling and simulations are terms for the used modelling parameters and the obtained simulation results. For the exchange of metadata information, one or several of the following metadata formats should be used if feasible:

- ISA-Tab: Investigation-Study-Assay, a Tab-separated format with data, model and Standard Operating Procedures (SOP's) stored under an ISA instance;
- ISA-JSON: Investigation-Study-Assay in JSON format;
- FHIR: Fast Healthcare Interoperability Resources, a collection of semantically richly annotated healthcare data.

A list of important general metadata standards is given in Table A.21 of Annex 2.

22.3.2 Minimum reporting guidelines for Clinical Decision Support Systems (CDSS)

Part 2 of ISO/TS 9491 describes **CDSSs based on AI models**. MINIMAR (Minimum Information for Medical AI Reporting) describes the **minimum set of required metadata**. It also describes recommendations for clinical studies involving computational models for personalised medicine. Other minimum reporting guidelines for a CDSS, are defined by the EQUATOR³³¹ (Enhancing the QUALity and Transparency Of health Research) network or by the TRIPOD³³² (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) consortium.

Examples are CONSORT (Consolidated Standards of Reporting Trials) for reporting of randomised trials, SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), a standard for trial protocols, STARD (STAndards for the Reporting of Diagnostic accuracy studies), DECIDE (Developmental and Exploratory Clinical Investigations of DEcision), CLAIM (Checklist for Artificial Intelligence in Medical Imaging), PRISMA (Preferred Reporting Items for Systematic Review and Meta-analysis) and PROBAST (Prediction model Risk Of Bias Assessment Tool).

There is a corresponding set of **guidelines for AI-based models and trials**. These guidelines have some AI-specific items added to the checklist, *e.g.*:

- CHEERS-AI: economic evaluations to estimate the cost effectiveness of AI interventions;
- CLAIM: for medical imaging studies using AI;
- CONSORT-AI: for clinical trials evaluating interventions with an AI component;
- DECIDE-AI: for CDSSs driven by AI;

³³⁰ Hucka *et al.* Front Bioeng Biotechnol 2015;3:19 . <https://doi.org/10.3389/fbioe.2015.00019>

³³¹ <https://www.equator-network.org>

³³² <https://www.tripod-statement.org>

- PRISMA-AI: Systematic reviews and meta-analyses of interventions;
- PROBAST-AI: for prediction model studies based on AI;
- SPIRIT-AI: for clinical trial protocols with an AI component;
- STARD-AI: for reporting of diagnostic accuracy studies based on AI;
- TRIPOD+AI: for diagnostic or prognostic prediction models based on AI.

Part 2 of ISO/TS 9491 also describes the **requirements and recommendations for SOPs** (Standard Operating Procedures), the data collection by using the Clinical Record Forms (CRFs), the data cleaning, pre-processing and post-processing steps to ensure the readiness of the data for use by the computational models and the data analytics pipeline with the steps study design, pre-processing, feature selection, model training and validation. Furthermore, it describes the design of integrated data repositories and the data modelling process.

22.3.3 Terminologies/ontologies for description and annotation of data and model (components)

Ontologies for Systems Biology / Systems Medicine / Systems Pharmacology simulation shall be used for the semantic annotation of algorithms and the dynamic behaviour of models where possible. Some ontologies are listed in the Annex B.3 of the ISO 20691³³³ document. Others are linked by ontology web portals like *e.g.*, the Open Lookup Service³³⁴ (OLS), BioPortal³³⁵, OboFoundry³³⁶ OntoBee³³⁷ or the semantic lookup platform³³⁸.

Ontologies typically used for *in silico* medicine modelling and simulation are listed below, with a more comprehensive list of ontologies is given in Table A.19 in Annex 2.

- Computational Neuroscience Ontology (CNO);
- An ontology to characterise differences in versions of computational biology Models (COMODI);
- Human Physiology Simulation Ontology containing concepts for simulations, models and algorithms (HUPSON);
- Kinetic Simulation Algorithm Ontology for describing simulation algorithms (KiSAO);
- Mathematical Modelling Ontology (MAMO);
- Ontology of Physics for Biology (OPB) containing physics concepts to describe the dynamics of biological systems;
- An ontology for pharmacokinetic models (PK ontology);
- Precision Medicine Ontology (PreMedOnto³³⁹);
- Systems Biology Ontology (SBO) with terms useful for describing computational modelling;
- Terminology for the Description of Dynamics (TEDDY), an Ontology for the description of control elements and dynamics in systems biology and synthetic biology;
- Unified Code for Units of Measure (UCUM), defined by the Regenstrief institute.

OpenMinds (Open Metadata Initiative for Neuroscience Data Structures) is a metadata framework for neuroscience graph databases defined by the Human Brain (HBP) and EBRAINS projects, containing metadata models, libraries of controlled terminologies, brain atlases and common coordinate spaces. Each such metadata model consists of modular metadata schemas defining descriptions of entities from neuroscience.

22.3.4 Clinical languages, terminologies and code systems

Clinical terminologies and code systems are mainly used for semantic annotation of Electronic Health Record (EHR) data and in FHIR³⁴⁰ resources. The clinical languages like Arden syntax and CQL allow the representation of medical knowledge in Clinical Decision Support Systems (CDSSs) and can be used as a basis for explainability components.

³³³ <https://www.iso.org/standard/68848.html>

³³⁴ <https://www.ebi.ac.uk/ols/index>

³³⁵ <https://bioportal.bioontology.org>

³³⁶ <https://obofoundry.org>

³³⁷ <https://ontobee.org>

³³⁸ <https://semanticlookup.zbmed.de>

³³⁹ <https://bioportal.bioontology.org/ontologies/PREMEDONTO>

³⁴⁰ <https://www.hl7.org/fhir/resource.html>

The main clinical languages and terminologies are listed below, with a more comprehensive list of clinical languages, terminologies, and code systems is given in Table A.20 in Annex 2 (further medical vocabularies are also listed in the Unified Medical Language System (UMLS)³⁴¹).

- International Classification of Diseases (ICD-11³⁴²);
- Logical Observation Identifiers Names and Codes (LOINC³⁴³) for reporting laboratory test results;
- Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT³⁴⁴).

22.4 Semantic annotation and taxonomy for the VHT

Metadata will be implemented for the uplift of the information coming from the data. Each data object will be accompanied by **well-defined sets of metadata**.

- Organisational paradigm metadata – they are associated to the six dimensions to describe the data object (both DOT and DOP). Such metadata are essential for the uplift of the information for that type of data
- Functional metadata – minimum metadata for the VHT being actionable.
- Accessibility metadata – they will give information about the “appropriateness of use” from different point of views:
 - Legal – either conditions for which the person from which that data has been gathered gave his or her consent or description of the other legal basis for the data use;
 - Ethical – approval number from the ethical committee;
 - Terms of use – terms under which the resource can be used (“civil code of supplementary laws”) given by the user who upload the resource;
 - Accessibility – how the data object can be accessed;
 - Credibility – information about trueness and precision;
- Optional metadata – metadata that could be added to better describe the data object, but not essential for the implementation of the VHT (*i.e.*, atlas positioning through the six dimensions and actionability).

There are also two **metadata transversal categories**.

- Necessarily standardised metadata - Metadata to make the infrastructure work, for which the content must necessarily be standardised.
- Optional standardised metadata - Metadata for which a precise standard is not necessary.

The metadata belonging to the first category are those essential for processes automatization, essential either for the infrastructure or for describing the data object. The minimum set of annotations will be decided by the community. Semantic mediation methods will be implemented to foster the upload of data objects also when coded with different standards. However, since the semantic mediation does not allow for a perfect mapping from a standard to another one, a human control will be added at the end of the procedure to check the quality of the mapping. This applies for publication purposes on the catalogue. If the data object has to be used in the platform for simulation, it must be coded with the standard adopted within the VHT (chosen by the community). Another possible approach is the Common Data Model – the header is pre-defined, the user can then decide how to code the information.

22.5 Missing standards still under development

Some desirable standards are still lacking, and the time frame of the EDITH project was too short for the full definition of needed new standards. Nevertheless, the definition process for the following standardization projects were already started and will continue in the years to come:

- **Adaptation of ASME VV-40:2018 standard:** The goal is the adaptation of the ASME VV-40 medical device standard for the risk-informed credibility assessment of biochemical models (SaMD). The new standard should include the assessment of agent-based models as well as data-driven (AI/ML) models. A joint IEC TC 62 / ISO TC 276 working group has been established.
- **Standardization for agent-based modeling: openVT (Virtual Tissue):** The goal is to develop a multi-scale agent-based standard with abstract concepts (“objects”, “processes”, “connections”), which subsumes the current MorpheusML, MultiCellIDS, MultiCell3D standards and the proprietary CompuCell3D format. It is planned to also define a proper uniform terminology and that the new standard supports the SBML, Antimony, CellML, MaBoSS and the BioNetGen network standards. This standard is currently under discussion in the COMBINE network.

³⁴¹ <https://www.nlm.nih.gov/research/umls/sourcereleasedocs/index.html>

³⁴² <https://icd.org>

³⁴³ <https://loinc.org>

³⁴⁴ <https://www.snomed.org>

- **Development of a new workflow standard for the DTs** as decided in the break-out session of the Paris meeting. Until now no official definition group was established, since it should first be analyzed and made explicit, which features and concepts are missing in the current CWL (Common Workflow Language) standard and what are the expected additional features from the new standard.
- **Standard for the semantic description of DTs in healthcare:** There are already plenty of official standards from ISO, IEC, IEEE, ITU and W3C available describing in detail the technology of important components and parts (virtual entities, physical entities, data, connection, services, sensors, ...) of technical digital twins³⁴⁵, but the overall architecture of a digital twin solutions is still under debate and keeps going. In addition, there are the DTDL³⁴⁶ (Digital Twin Definition Language) defined by Microsoft and the W3C Web of Things (WoT) standard³⁴⁷.
For digital twins in healthcare, one needs a specific standard for their semantic description. Such a standard is currently under discussion in the COMBINE network and comprises several levels of metadata (biological, technical, legal and descriptive metadata) describing a DT³⁴⁸ (Table 3).

Table 3: Biological, technical, legal and descriptive metadata describing a DT

Technical metadata:	Biological metadata:
<ul style="list-style-type: none"> • components of a DT (data, model or ensemble of models, comparator: divergence between the predicted state and observed state, actuator: does something in response to comparator output, control: decision what to do, <i>i.e.</i>, model reacts to events) • DT interface: input and output data (data type, data format, frequency of data acquisition, availability, reliability, security, units) • required runtime parameters (significance level,...) • parametric variability (how to adapt a generic model to become an individual, <i>i.e.</i> a personalized DT?) • “Parameter cloud” within which the model functions correctly • technical integration into multi-scale/multi-organ models (VHT as a collection of DTHs): modularization → need for internal interfaces and a modularized architecture • model validation (ASME VV-40, identifiability, observability, reproducibility, explainability, controllability). Problem for personalized models: validation can only be done retrospectively! • language runtime (C, CLR, JRE, R, Julia, Jupyter, Mathematica, Matlab/GNU Octave, Python) • runtime dependencies (libraries, operating system) • execution hardware (workstation, HPC cluster, cloud, quantum) • execution environment (containerized, WfExS–Workflow Execution Service, web application running in the browser, command-line) • container type (Docker, Apptainer) • execution/solver software (COPASI, libroadrunner, Morpheus, Tellurium, Vivarium, XPPAUT) 	<ul style="list-style-type: none"> • organism (human for the VHT) • domain of model (medical, epidemiological, behavioural, ecological, populational) • sort of model (ensemble model, single model) • model types (physiological / pathophysiological, biomechanical, imaging, systems biology / systems medicine / systems pharmacology / systems toxicology) • purpose (research, monitoring, diagnostic, prognostic, therapy choice, therapy optimization, risk prediction, education, biomanufacturing, healthcare/hospital management) • system level (molecular, pathways, subcellular, cell, tissue, organ, organ system, whole body, communities (epidemiological or population statistical models)) • biological integration into multiscale models (cell-cell signalling (tissue factors), nervous signals, hormones, microbiome interactions (skin, gut), PBPK (Physiology-Based Pharmacokinetic model)) • information/data basis of model (mechanistic model, data-based (statistical) model, AI/ML model, hybrid model, agent-based model) • target value of prediction (scale value, survival time, yes/no, risk, recommendation)
	Legal metadata:
	<ul style="list-style-type: none"> • privacy/data protection issues (anonymization, pseudonymization, consent, GDPR) • data security • data encryption • Intellectual Property Rights (IPR) • license information: list from the Open Software Initiative or Software Package Data Exchange (SPDX)
	Descriptive metadata:
	<ul style="list-style-type: none"> • title of the DT • description of the DT

³⁴⁵ <https://digitaltwin1.org/articles/2-4>

³⁴⁶ <https://azure.github.io/opendigitaltwins-dtdl/DTDL/v3/DTDL.v3.html>

³⁴⁷ <https://www.w3.org/TR/wot-thing-description11>

³⁴⁸ <https://zenodo.org/records/14198150>

<ul style="list-style-type: none"> • model type (ABM, algebraic equations, Bayesian, Boolean, Graphical, linear equations, metabolic network, ODE, PDE, Petri, stoichiometric) • model format (CellML, FieldML, SBML, R, XPP) • data format (XML, RDF, JSON, JSON-LD, YAML, FHIR, Apache Avro, Google protobuf) for patient / healthcare data files • credibility (from no credibility assessment done, to certified by regulatory agency for a specific CoU) • clustering (from individual to entire population) 	<ul style="list-style-type: none"> • targeted group (age, sex, other inclusion/exclusion criteria) • regulatory aspects (approval by FDA and/or EMA; “approval number”) • links (to publication, documentation website, source code) • version • last update date • funding • authors
--	--

23 Regulatory science

23.1 General context and regulatory actors

23.1.1 The intended use of DTs

The regulatory process for Digital Twins depends on their intended use and end user³⁴⁹. **Personal health monitoring and forecasting** technologies can empower patients and healthcare providers in their health and care journey for better rehabilitation and management of health conditions. The primary users are patients, healthcare providers in partnership with patients. In **Clinical Decision Support Systems** (CDSS), DTs assist and support healthcare professionals with informing, guiding, diagnosing, predicting and making decisions about treatments. The primary users are clinicians and patients in partnership with clinicians. Finally, when applied in the context of *in silico* trials for the design, development and assessment of biomedical products, DTs augment physical testing and trials. *In silico* trials simulate the potential effects and outcomes of medical products (drugs, devices, advanced therapies) in order to evaluate their safety, efficacy and effectiveness. Unlike conventional trials that use physical experiments and observations, *in silico* trials use virtual or digital methodologies, often referred to *in silico* trials as ‘methodologies’ or ‘solutions’. The primary users of this application domain are product developers, regulators and health technology assessors.

23.1.2 Regulatory actors and guidance

Given that the focus in this roadmap is on Europe, this section does not provide an exhaustive overview of the regulatory actors across the globe. We will focus on the European actors, along with the those in the US. A review of the regulatory documents relevant to the *in silico* methodologies used to evaluate new medical products across the globe is provided in the Annex of the GSP book³⁵⁰.

Regulatory agencies play an essential role in advancing the safe and effective use of computational modelling and simulation in healthcare by providing guidance, setting standards, and promoting innovation. The **U.S. Food and Drug Administration** (FDA) has been proactive in exploring DT model technologies, both knowledge-driven and data-driven, issuing key guidance documents such as the 2016 guideline on reporting computational modelling studies in medical device submissions³⁵¹, the 2018 guidance on PBPK models³⁵², a draft framework for assessing model credibility that aligns with the ASME VV-40 standard³⁵³, and recently, a draft guidance on the use of Artificial Intelligence to support regulatory decision-making for drug and biological products.

Similarly, the **European Medicines Agency** (EMA) supports the integration of these methodologies through initiatives like the "Innovation Task Force" and guidance documents, including the guideline on reporting PBPK models³⁵⁴ and the guidance for qualifying novel methodologies in medicine development³⁵⁵. EMA is actively involved in a number of **initiatives and networks that are directly relevant for the VHT**. One example is the Big Data Initiative³⁵⁶ to enable use of Real-World Data and facilitate its integration in regulatory decision making. A related example is the DARWIN EU® Data Analysis and Real World Interrogation Network³⁵⁷. DARWIN-EU also serves as a use case in HealthData@EU³⁵⁸ where it operates as a node in the EHDS2 pilot for the secondary use of health care data. A last example is the recent publication of a reflection paper on the use of AI in the medicinal

³⁴⁹ Pappalardo F, Wilkinson J, Busquet F, et al. Toward A Regulatory Pathway for the Use of *in silico* Trials in the CE Marking of Medical Devices. *IEEE J Biomed Health Inform.* 2022;26(11):5282-5286. doi:10.1109/JBHI.2022.3198145

³⁵⁰ <https://link.springer.com/book/10.1007/978-3-031-48284-7>

³⁵¹ <https://www.fda.gov/media/87586/download>

³⁵² <https://www.fda.gov/media/101469/download>

³⁵³ <https://www.fda.gov/media/154985/download>

³⁵⁴ https://www.ema.europa.eu/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpb-modelling-simulation_en.pdf

³⁵⁵ https://www.ema.europa.eu/documents/regulatory-procedural-guideline/qualification-novel-methodologies-drug-development-guidance-applicants_en.pdf

³⁵⁶ <https://www.ema.europa.eu/en/about-us/how-we-work/big-data>

³⁵⁷ <https://www.ema.europa.eu/en/about-us/how-we-work/big-data/real-world-evidence/data-analysis-real-world-interrogation-network-darwin-eu>

³⁵⁸ <https://ehds2pilot.eu/>

product lifecycle. EMA's **regulatory role for medical devices** is rather limited to the assessment of certain categories of medical devices and *in vitro* diagnostics and combination products. Additionally, EMA provides scientific opinions for certain high-risk medical devices through specific expert panels who benefit from EMA's technical and scientific support³⁵⁹.

Notified bodies are commercial organizations designated by an EU country³⁶⁰ (and audited by their National Competent Authorities) to **assess the conformity** of certain products before being placed on the market. These bodies carry out tasks related to conformity assessment procedures set out in the applicable legislation, when a third party is required. They are assigned an important role in various health related regulations, such as the Medical Devices Regulation, the *In vitro* Diagnostic Medical Devices Regulation and the AI Act.

23.2 Current status of regulatory pathways for DT models

23.2.1 Standards as drivers for regulatory science

In 2016, the FDA Centre for Devices and Radiological Health (CDRH) published a guideline³⁶¹ entitled “*Reporting of Computational Modeling Studies in Medical Device Submissions*”, followed in 2018 by the publication of the **ASME VV-40:2018** technical standard, entitled “*Assessing Credibility Of Computational Modelling Through Verification And Validation: Application To Medical Devices*”. This opened the possibility of producing evidence of safety or efficacy for new medical devices in their regulatory process. Soon the credibility assessment approach proposed by the VV-40 was also recognised by the FDA CDRH as valid for the certification of SaMD with predictive capabilities. In the European Union, notified bodies have already accepted *in silico* evidence to reduce *in vitro* testing. However, the recent start of standardisation activities by ISO/IEC aimed to develop standards with a scope similar to the VV-40 (discussed in the previous chapter), opens up a future where also the CE marking of medical devices based on evidence provided by *in silico* methodologies, or the CE marking of predictive SaMD, will be possible using a **harmonised standard**. And meanwhile, it is reasonable to expect the EU notified bodies to accept transitory evidence of credibility based on the ASME VV-40:2018.

23.2.2 Key regulatory pathways

Two key regulatory pathways have been established for the adoption of *in silico* methodologies and digital twins in medical device and drug development (discussed in more detail in the next section):

Certification of Software as a Medical Device (SaMD): regulatory authorities such as the FDA's Center for Devices and Radiological Health (CDRH, 510(k)) and the EU's CE-marking process recognize SaMD or Medical Device Software (MDSW) as a distinct category of medical devices. These pathways are particularly focused on predictive-capability SaMDs, exemplified by solutions like HeartFlow, which enhance decision-making in clinical settings.

Qualification of *in silico* methodologies: both the FDA and EMA offer pathways to qualify tools used in medical device and drug development. While the FDA provides qualification pathways for both devices and drugs, the EMA currently limits its qualifications to drug development tools. Although not mandatory, qualification is recommended and involves seeking advice from regulatory bodies and submitting a formal request for a qualification opinion. A positive opinion allows developers to use the methodology for generating evidence in marketing authorization applications.

In addition to these pathways, **Digital Health Technologies (DHT)** are starting to be considered by regulators in the US and EU, where the key interest seems to revolve around embracing digital tools to capture clinical trial data, for which regulatory frameworks are yet to evolve³⁶². As indicated by FDA, DHTs include electronic sensors, computing platforms and information technology. Portable DHTs that

³⁵⁹ Role of EMA on Medical devices - <https://www.ema.europa.eu/en/human-regulatory-overview/medical-devices>

³⁶⁰ <https://webgate.ec.europa.eu/single-market-compliance-space/notified-bodies>

³⁶¹ <https://www.fda.gov/media/87586/download>

³⁶² [FDA's Framework for the use of digital health technologies in drug and biologics product development](https://www.fda.gov/media/166396/download?attachment) - <https://www.fda.gov/media/166396/download?attachment>

may be worn, implanted, ingested, or placed in the environment allow real-time and continuous collection of data from trial participants in their homes or at locations remote from clinical trial sites.

23.2.3 Successful examples

Despite significant hurdles, *in silico* medical device solutions are already making a positive impact on patients by piloting through the regulatory pathway for medical devices. Here, we discuss three successful examples of *in silico* medical devices navigating regulatory and reimbursement landscapes across various regions³⁶³.

- The example of **HeartFlow**³⁶⁴ was discussed previously in this roadmap. Heartflow Inc. is a USA-based SME, established in 2010 (Mountain View, CA). They focus on coronary artery disease and have several products on the market, including the Fractional Flow Reserve Computed Tomography (FFRCT®, generations 2014, 2018, and 2021), Roadmap™, and HeartFlow ONE™. HeartFlow's non-invasive HeartFlow FFRCT Analysis creates a personalized three-dimensional model of the heart. Clinicians use this model to evaluate the impact a blockage has on blood flow and determine the best treatment for patients. It is certified as a SaMD, Class IIa (EU, 2011) and Class 2 deNovo (510k) (US, 2014). In addition, it has obtained regulatory approval in Japan (2015) and Canada (2019). A positive reimbursement decision was obtained in the USA, UK and Japan (2018). The UK HTA body NICE estimated that by 2021/22, HeartFlow FFRCT® may lead to **cost-savings of £391 per patient** (2021) and will potentially be used in 40,000 patients annually, saving the UK National Health Services (NHS) at least £9.1 million, see details³⁶⁵.
- **FEOps**³⁶⁶, a Materialise company, is a spin-off from the University of Ghent and was incorporated in 2009. It has two products on the market: the TAVIguide™ since 2015, and the HeartGuide™ since 2019 (update 2023). They focus on the challenge of Left Atrial Appendage Occlusion (LAAO) and are recognized pioneer in the field of physics-based simulations for minimally invasive cardiovascular devices and procedures. HeartGuide™ is certified as a SaMD, Class IIa (EU, 2015) and Class 2 (USA, 2022). Regulatory approval was also granted in Japan (2019) and New Zealand (2020). A positive reimbursement decision was obtained in the USA in 2024.
- **inHeart**³⁶⁷, a French SME was established in 2017 in Pessac (France). In 2019, it launched their product inHEART™. The product focusses on cardiovascular – Electrophysiology interventions, delivering an AI-enabled, digital twin of the heart to advance the care of patients living with cardiac disease, *e.g.* the digital twin of the heart for image guided ablations. inHeart™ is certified as a SaMD, Class IIa (EU, 2019) and Class 1 and then 2 (USA, 2022). A positive reimbursement decision was obtained in the USA in 2024.

23.3 DTs for clinical decision support and personal health forecasting

The U.S. regulatory framework is based on guidance³⁶⁸ from the International Medical Device Regulators Forum (**IMDRF**), a global body established to harmonise medical device regulations. The IMDRF defines SaMD as software intended for one or more medical purposes that perform these purposes independently of a hardware medical device. SaMD can also serve as a personal health forecasting tool when designed for use by patients or caregivers. These tools must comply with regulatory pathways based on risk classification, such as Premarket Notification (510(k)), De Novo Classification Request and Premarket Approval (PMA). Key steps for bringing SaMD to market (in the USA) include classifying the device, preparing a premarket submission, sending it to FDA, interacting with FDA for review and, finally, complying with the applicable regulatory controls.

Another important guidance published by the **FDA** deals with how to collect and report clinical evidence for software as medical devices³⁶⁹. The typical clinical evaluation process is divided into three steps. The first step is the valid clinical association step, which aims to verify that the association between the SaMD output and the targeted clinical condition is supported by evidence. The second step is the analytical validation step, which aims to generate evidence that shows that the output of your SaMD is technically what you expected. The third step is the clinical validation step, which aims to

³⁶³ Diabeloop DBLG1 - <https://www.diabeloop.com/>

³⁶⁴ HeartFlow™ (U.S.A) - <https://www.heartflow.com/heartflow-ffrct-analysis/>

³⁶⁵ <https://www.nice.org.uk/guidance/mtg32/resources/heartflow-ffrct-for-estimating-fractional-flow-reserve-from-coronary-ct-angiography-pdf-64371991952581> (published 2017, updated 2021)

³⁶⁶ FEops Heartguide™ <https://www.feops.com/src/Frontend/Files/MediaLibrary/09/press-release-feops-march-2019-final.pdf>

³⁶⁷ <https://www.inheartmedical.com/news/inheart-receives-ce-certification-under-new-mdr-for-novel-ai-based-digital-twin-of-the-heart>

³⁶⁸ <https://www.imdrf.org/working-groups/software-medical-device-samd>

³⁶⁹ <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/software-medical-device-samd-clinical-evaluation>

generate evidence that shows your SaMD has been tested in your target population and for your intended use, and that users can achieve clinically meaningful outcomes.

In the EU, Digital Twin technologies are regulated as software as a medical device under the European **Medical Device Regulation (MDR)**, effective since May 26, 2021 and discussed in detail in the next chapter. Certification is granted by notified bodies. Upon approval, devices receive a CE mark, indicating conformity. The MDR's introduction has placed significant strain on notified bodies due to the requirement to recertify all existing medical devices. This has limited their capacity to address innovative products, presenting challenges for the approval of cutting-edge technologies.

23.4 DTs for *in silico* trials

The FDA has published two guidelines, one for Medical Device **Development Tools (MDDTs)**³⁷⁰ and another for the qualification process for Drug Development Tools (DDTs)³⁷¹. MDDTs can be categorised into three types: non-clinical assessment models, biomarker tests, and clinical outcome assessments. We should note that among the non-clinical assessment models, we have physicochemical or biological-based computer models. Drug Development Tools (DDTs) are methods, materials, or measures that have the potential to facilitate drug development. Examples of DDTs may include but are not limited to, a biomarker used for clinical trial enrichment, clinical outcome assessment tools used to evaluate clinical benefit, and animal models. Once qualified, DDTs will be publicly available for use in any drug development program for the qualified context of use.

As mentioned above, an important document is the guidance published by the FDA in November 2023 to assess the credibility of computational modelling and simulation in medical device submissions. This document provides a framework that manufacturers can use to demonstrate that computational modelling and simulation (CM&S) models supporting regulatory submissions are credible. The guidance applies to CM&S models that are physics-based or knowledge-based and does not apply to standalone ML or AI-based models.

The FDA recognises the standard ASME VV-40 "*Assessing Credibility of Computational Modelling through verification, validation, and uncertainty quantification (VVUQ) activities*" published by the American Society of Mechanical Engineers (ASME) in 2018, and utilises the same key concepts. However, it provides a more general framework for demonstrating CM&S credibility in medical device regulatory submissions, incorporating different categories of credibility evidence.

As already mentioned, in the EU, the European Medicines Agency (EMA) solely regulates medical products and has published guidelines on the **qualification process of novel methodologies for medicine development**. EMA's Committee for Medicinal Products for Human Use (CHMP) first provides scientific advice to support the qualification of innovative development methods for a specific intended use in the context of research and development into pharmaceuticals, based on recommendations by the Scientific Advice Working Party. Following this initial phase, applicants must submit a document – referred to as a briefing book – that describes the novel methodology and includes a list of questions and applicant positions.

EMA typically first offers advice on protocols and methods intended for the development of a novel method. This advice is based on the evaluation of the scientific rationale and preliminary data submitted by the applicant. Based on qualification advice, EMA may propose a letter of support as an option when the novel methodology under evaluation cannot yet be qualified but shows promise based on preliminary data. EMA may also issue an opinion on the acceptability of a specific use of a method, such as the use of a novel methodology or an imaging method in the context of research and development. Before the adoption of the methodology, EMA makes its evaluation open for public consultation by the scientific community. This ensures that the EMA shares information, as agreed with the applicant, and is open to scientific scrutiny and discussion.

³⁷⁰ <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/qualification-medical-device-development-tools>

³⁷¹ <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/qualification-process-drug-development-tools-guidance-industry-and-fda-staff>

Additionally, even though this standard is related to applications for medical devices and not medicinal products, the EMA suggests the use of the ASME VV-40 standard to provide evidence about the credibility of computational models used in Drug Development Tools (DDT).

As previously mentioned, no central authority exists in Europe for medical devices, and no European guidelines exist yet for the qualification of Computational Modelling and Simulation (CM&S) as Medical Device Development Tools (MDDT). However, the **MDR explicitly mentions modelling and simulation**: “Where appropriate, the results of biophysical or modelling research, the validity of which has been demonstrated beforehand, may be considered in relation to the device requirements regarding design and manufacture.”

Therefore, nothing in the current EU regulation would prevent companies from producing regulatory evidence to support the marketing authorisation of a medical device obtained through *in silico* methods. Some orthopaedic companies already use finite element model results to support their request for the CE marking of new orthopaedic devices. Because no harmonised standard is available, they sometimes follow the ASTM F2996-20 “Standard Practice for Finite Element Analysis (FEA) of Non-Modular Metallic Orthopedic Hip Femoral Stems” to support the credibility of this *in silico* evidence.

A tangible outcome of the efforts to promote this debate within the regulatory community is the recent publication of the revised technical standard ISO 21535:2023, “Non-active surgical implants. Joint replacement implants. Specific requirements for hip-joint replacement implants.” The standard states: “Theoretical analysis and modelling, including finite element analysis, can also be used to select the most appropriate size(s) of component(s) for testing the worst case(s) (e.g., see ASTM F2996). If used, the credibility of such modelling for its context of use shall be demonstrated (see ASME V&V 40-2018).”

This marks an important precedent, extending beyond orthopaedic companies. By formally recognising ASME V&V 40 as the benchmark for assessing the credibility of predictive models, the ISO standard opens the door for *in silico* evidence to be used more broadly in regulatory processes.

23.5 Regulatory science for the VHT

Translating VHT technologies into the clinic is one of the key desired outcomes of the VHT initiative. The combination of novelty and complexity of the VHT technologies, with the fragmentation of the regulatory decision-making landscape, and the absence of established standards and precedents for evaluating their safety and efficacy in healthcare, **could slow down the translation of VHT technologies** into the clinic. In addition, the lack of sufficient, high-quality data for VHT development limits the ability to develop comprehensive and credible models that can be effectively used for research, regulatory assessment, clinical applications, and commercialization.

The VHT Infrastructure is an essential element in reaching the goal of clinical translation. The infrastructure should not only facilitate the pooling of resources, but also the **development and credibility assessment** of the DTs. Throughout this roadmap, the credibility of VHT resources has been discussed from many different angles, including conceptually, technically, infrastructurally, from a standards perspective, and, in the next chapters, also legally, ethically and socially.

One clear way in which the VHT Infrastructure can help the credibility assessment of DTs by offering specific services, workflows and resources. These **services and workflows** could range from providing assistance in the uptake of the correct standards, to assisting users in performing credibility assessment, following the VVUQ principles. An example of such a dedicated workflow is the VECMA toolkit³⁷², that includes dedicated workflows for uncertainty quantification on HPC. FDA’s regulatory science tools³⁷³ have been developed in this spirit, by adopting tools developed/checked by the regulator, the entire regulatory process can be accelerated. **Validation data sets** can be collected in the Catalogues and stored in the Repository (in case the data is not yet stored elsewhere). These data sets can come from (anonymized) clinical and real-world data³⁷⁴, or can be generated in the Infrastructure (synthetic data). Finally, establishing federation with digital regulatory sandboxes, such as the one under development in the REALM project can speed up the regulatory acceptance.

³⁷² <https://www.vecma-toolkit.eu/>

³⁷³ <https://www.fda.gov/medical-devices/science-and-research-medical-devices/catalog-regulatory-science-tools-help-assess-new-medical-devices>

³⁷⁴ See for instance the Validation Collection of the InSilicoWorld project: <https://zenodo.org/communities/insilicoworld-openaccess/records?q=validation%20collection&l=list&p=1&s=10&sort=bestmatch>

23.6 Digital Twin use cases

Box 39: Use Case - Glycemic control in ICU patients (regulatory).

Use Case: *Glycemic control in ICU patients*

Status: *start-up company, solution in clinical trial*

Website: www.insilicare.com

InSiliCare's AI-powered Digital Twin is a clinical decision support solution for ICU patients requiring glycemic control. It provides personalized, safe, and effective management of blood glucose levels and nutrition delivery. Insulin and nutrition treatments are calculated to maximise safety from hypoglycemia, while controlling patient blood glucose levels and optimizing nutrition towards a configurable physician-determined practice of care.

Insilicare, a spin-off company (www.insilicare.com) has been created to commercialize this solution. To date, Insilicare is finalizing the QMS and technical documentation to comply with the MDR requirements. The conformity assessment has not started yet.

Box 40: Use Case - Osteoporotic fracture risk prediction (regulatory).

Use Case: *Bologna Biomechanical Computed Tomography for osteoporotic fracture risk prediction*

Status: *clinical research studies*

Bologna Biomechanical Computed Tomography (BBCT) is a Digital Twin methodology designed to predict the mechanical strength of the femur under critical loading conditions in osteoporotic patients. Quantitative Computed Tomography (QCT) scans of the hip region and patient data inform a subject-specific Finite Element (FE) model able to predict the risk of hip fracture at the time the CT is performed (ARF0).

Model credibility assessment was performed following the ASME VV40 guidelines for the use of BBCT to determine the optimal effective dose for a new anti-osteoporosis drug in adults and older adults (from 55 years) according to multi-dose Phase II studies³⁷⁵. BBCT itself is currently not certified for clinical use and is limited to clinical research studies³⁷⁶. To demonstrate the validity of the methodology, it could be valuable to perform power calculations to compare the differential discriminatory power of ARF0 against BMD data. In a future, ideal version of BBCT (at TRL8+), the solution would function as a decision-support system, providing individualised predictions of hip fracture risk. This solution would fall into the category of medical devices known as Software as a Medical Device (SaMD). Achieving compliance with medical device regulations would require obtaining appropriate certifications, such as the CE mark in Europe or FDA clearance in the United States. Under the European Union's Medical Device Regulation (MDR), the software would likely be classified as Class IIa, given its potential impact on patient care decisions. A clinical evaluation process would need to be undertaken to assess the safety and performance of the software in a clinical setting. This process typically involves gathering clinical data, conducting studies, and generating evidence to demonstrate the software's effectiveness in achieving its intended purpose.

Box 41: Use Case - Universal Immune System Simulator for Tuberculosis (regulatory).

Use Case : *Universal Immune System Simulator for Tuberculosis*

Status: *augmenting clinical trials*

Website: <https://www.mimesis.srl/>

The Universal Immune System Simulator (UISS) offers a robust, clinically validated platform to address these challenges by simulating and predicting immune responses *in silico*. UISS-TB has demonstrated its utility in optimizing therapeutic strategies for Tuberculosis, evaluating drug regimens, and predicting outcomes.

UISS-TB is being developed as an innovative computational platform to support *in silico* trials and model-informed drug development (MIDD), focusing on advancing regulatory science and clinical applications through EMA guidance.

The EMA acknowledges the platform's potential to predict circulating interferon gamma (IFN- γ) changes over time as a response to treatment dose in virtual patient cohorts. This feature is particularly valuable for informing dose selection in phase IIa trials of therapeutic vaccines targeting latent pulmonary TB. However, the EMA

³⁷⁵ <http://doi.org/10.1016/j.cmpb.2023.107727>

³⁷⁶ <https://zenodo.org/records/10948478>

emphasizes the need for further validation of IFN- γ as a biomarker and encourages expanding clinical datasets to support broader regulatory acceptance.

Key elements of UISS-TB's regulatory strategy include³⁷⁷:

1. **Robust Validation:** The platform's development adheres to ASME VV-40:2018 standards for verification and validation, ensuring rigorous technical and credibility assessments.
2. **Uncertainty Quantification:** A comprehensive Monte Carlo-based approach evaluates how input uncertainties propagate through the model, providing transparent and reproducible outputs.
3. **Regulatory Engagement:** UISS-TB builds on established precedents, such as the EMA's qualification of the hollow fibre system for TB, while addressing its unique characteristics as an agent-based modelling platform.

UISS-TB has received a Letter of Support from EMA³⁷⁸, in which the agency underscores the importance of UISS platform's advanced, comprehensive solutions in TB therapeutic vaccine development.

Box 42: Use Case - Epileptogenic zone localisation for surgical planning in epilepsy patients (regulatory).

Use Case : Epileptogenic zone localisation for surgical planning in epilepsy patients

Status: in clinical use

Website: <https://www.cloudsofcare.com/>

Persyst ESI powered by Epilog is a neuroimaging solution that automatically combines scalp EEG data with a patient's MRI to perform Electrical Source Imaging (ESI). It pinpoints the origin of brain activity linked to seizures, helping clinicians accurately localize the epileptogenic zone—critical for surgical planning in epilepsy patients.

Persyst ESI powered by Epilog³⁷⁹ is an FDA class II approved device and is MDR Class IIA certified.

Box 43: Use Case - the Atrial Modelling Toolkit for cardiovascular Digital Twins (regulatory).

Use case : the Atrial Modelling Toolkit for cardiovascular Digital Twins

Status: for research purposes

Website: <https://github.com/pcmlab/atrialmtk>

The Atrial Modelling Toolkit³⁸⁰ (**atrialmtk**) aims to overcome the challenges of constructing cardiac models at scale through the development of a robust, open-source pipeline for bilayer and volumetric meshes for atrial models. Atrialmtk is a research tool not (yet) intended for clinical use.

Following best practices are already implemented related to AI and data:

- European Society of Cardiovascular Radiology best practices for artificial intelligence³⁸¹
- Reproducibility check of modelling pipeline for making left atrial models from MRI data³⁸².
- Data inputs/flow/outputs: developed pipeline with multiple input options to account for both standard clinical imaging format & anatomies from artificial datasets *e.g.* those generated by statistical shape models, possible different operating systems of the user, possibility to integrate & run on HPC facilities as well as local laptop/desktop PC etc.
- Standard way of visualising/quantifying simulation outputs *e.g.* common in other modelling literature or in clinical data

For the example of *in silico* clinical trials, following best practices are adhered to:

- Best practice for reporting *in silico* clinical trials, including verification, validation and uncertainty quantification³⁸³.
- Patient demographics are reported and codes are made available open source on github.
- Validation studies show that the models are calibrated to sinus rhythm cardiac pacing data, and that simulation metrics agree with clinical metrics (atrial fibrillation rates, recurrence rates and therapy

³⁷⁷ <https://zenodo.org/records/10948478>

³⁷⁸ https://www.ema.europa.eu/en/documents/other/letter-support-universal-immune-system-simulator-tuberculosis-disease-model-uiss-tb-dr_en.pdf

³⁷⁹ <https://www.persyst.com/technology/electrical-source-imaging/>

³⁸⁰ <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2023.0038>

³⁸¹ Weikert et al. European Radiology. 2021; 31:3909–3922

³⁸² <https://pubmed.ncbi.nlm.nih.gov/37301099/>

³⁸³ <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1012289>

outcomes). Uncertainty is investigated by simulating the effects of uncertainty in measurements used for calibration on model outcomes³⁸⁴.

When transferring to a clinical decision support tool for an individual with atrial fibrillation, certification is required *e.g.* CE mark (UK, Eu) NICE and MHRA (UK), Class IIa – due to potential impact on clinical decisions. If developed further for clinical use then **atrialmtk** would need to meet ISO/regulatory standards for medical device/software.

³⁸⁴ <https://pubmed.ncbi.nlm.nih.gov/36195766/> and <https://pubmed.ncbi.nlm.nih.gov/30969911/>

24 Health technology assessment and payers

24.1 General considerations

Health technology assessment (HTA) is a multidisciplinary process that uses systematic and explicit methods to evaluate the properties and effects of a health technology³⁸⁵. It can be used for multiple purposes, but the most common is to **evaluate the cost-effectiveness** of new treatments of instrumentation, and to support reimbursability evaluations.

Several **HTA assessments of Digital Twin technologies** are starting to appear³⁸⁶. The HTA assessment of a Digital Twin does not pose particular challenges; the HTA framework is general enough to be used for any technology, whether predictive or not. More interesting is the use *in silico* methodologies for the HTA assessment of new medical products, whether drugs or devices. The **use of computer modelling and simulation in Health Technology Assessment** could offer several advantages. *In silico* methods can be a cost-effective and time-efficient approach compared to conducting real-world trials or traditional clinical experiments. Also, computer models can estimate parameters that are difficult or impossible to measure and can predict long-term outcomes, enabling the assessment of potential impacts and benefits of new medical products over extended periods. So far, to the authors' knowledge, no concrete use of *in silico* methodologies has been reported to support HTA decisions.

While the use of *in silico* methodologies seems promising for improving HTA processes, it is essential to **examine its limitations and potential risks**. Considering the barriers that need to be addressed for the VHT to realise its full potential in advancing human health (PART 1), the ones that more closely interests the application on HTA and Payers are mostly related to data availability and quality, lack of standardisation, technical expertise and resources, validation and transparency. The **availability of high-quality data** is crucial to assess the validity and reliability of the modelling results. The lack of **standardised methods**, guidelines, or best practices to demonstrate the efficacy and cost-effectiveness of *in silico* methodologies in HTA can lead to inconsistencies in approaches, difficulties in interpreting and evaluating different models and results. Payers and HTA bodies may face barriers in terms of **expertise**, infrastructure, and funding necessary to utilise these models effectively. Also, introducing new methodologies and approaches, can face resistance due to cultural barriers and lack of familiarity especially when making time-sensitive decisions, which can limit the feasibility and practicality of using complex models. **Validating and transparently documenting** computer models and simulations is essential to assess their credibility. However, validating complex models can be challenging, and lack of transparency in model assumptions, algorithms, and data sources can hinder trust and acceptance among stakeholders, including payers.

The VHT can certainly support and facilitate these research activities aimed to explore the usefulness and the limits of using *in silico* methodologies for the HTA of medical products, but this requires dedicated (funding) incentives.

HTA actors play a crucial role of compiling information from various fields while also considering efficacy, safety, economic, and organizational aspects of health technology. This process is conducted in a systematic, transparent, unbiased, and rigorous manner. The purpose is to inform decision making by using explicit methods to promote an equitable, efficient, high-quality health system³⁸⁷. Whilst regulatory or market access licensing approval is mainly focused on the technical and safety profile of the medical device, HTA bodies have **different interests and, therefore, different evidence requirements**. Normally, the requirements aim to **inform policymakers** (and decision-makers in general) of the **rational allocation of resources within finite budgets** to the funding (or using) of healthcare interventions. For this reason, data required for market access might go beyond those used or developed for licensing, particularly in medical devices, where regulatory requirements have

³⁸⁵ https://en.wikipedia.org/wiki/Health_technology_assessment

³⁸⁶ <https://www.nice.org.uk/guidance/mtg32>

³⁸⁷ HTA Glossary. International Network of Agencies for Health Technology Assessment (INAHTA), Health Technology Assessment international (HTAi) and other partner organizations

historically been low. In addition to the classical **supply-driven reimbursement pathways** for VHT, where the VHT creators approach the payer community for coverage decisions, HTA actors involved in pre-commercial procurement of innovation can be key players in the strategic development of VHT through **demand-driven pathways and partnerships**. Interaction with HTA actors is crucial for the VHT, not only for the economic viability of VHT solutions, but also for the large-scale adoption of VHT technologies in the clinical practice, as reimbursement is a critical driver for penetration into current day value-based healthcare ecosystem.

24.2 HTA frameworks for prediction of DT technologies

While HTA bodies and their decision-making actors could play a crucial role in evaluating VHT, the necessary guidance and regulations to govern DT or VHT solutions are yet to be established. Nevertheless, there is an increasing trend to adapt existing HTA methodologies to assess the fast-emerging digital health technologies and clinical decision support systems, including AI-based digital medical devices. For instance, the HTA Core Model³⁸⁸ is seen as a valuable foundation, encompassing a wide range of topics for assessment, including legal, ethical, social, and organizational considerations³⁸⁹. This is further elaborated in the context of assessing AI-based health technologies, where challenges related to assessing self-learning AI models, the explainability of model decisions³⁹⁰, and interoperability³⁹¹ (which refers to how well an application integrates with other hospital systems) are assessed. To this end, a recent meta-analysis³⁹² of 31 policy documents and 9 academic publications, compiled by expert stakeholders focusing on AI-specific assessment elements, resulted in a framework comprising 29 assessment elements across four main domains: Technology & Performance, Human & Organizational, Legal & Ethical, and Transparency & Usability.

Considering the recent burst of regulatory approval of AI-based health technologies and the associated buzz around the potential of these evolutions reaching healthcare systems, HTA of AI-based technologies is gaining urgency and traction amongst the payer communities. While the necessary HTA frameworks to assess these AI technologies are fast emerging, a related sub-domain of **digital medical devices** (DMD)/digital health technologies are already anticipating and accommodating the assessment of VHT solutions (including those that use AI technologies). The HTA of DMD is, in many ways, a first step into the direction of the DTs and the envisioned VHT solutions. On this note, the following section explores the context of assessing DMD, in order to understand possible overlaps and relevance with respect to the reimbursement of Digital Twins in healthcare, within EU.

24.3 HTA frameworks and categorization of DT for reimbursement decisions within EU

24.3.1 Digital Health Technologies

Patient or clinician focussed Digital Twins in healthcare and VHT ecosystem solutions are likely to fall under **digital health technologies** that encompass **digital therapeutics, clinical decision support systems and personal health forecasts**, where they enhance the 4P principles in healthcare: predictive, preventive, personalized and participatory. Over time, the individual DTs catering to some of these problem areas are bound to evolve into connected or cascaded Digital Twins in healthcare. Eventually, the VHT solutions with integrated Digital Twins, are set to break pathology silos and offer integrated care pathways across the human pathophysiology.

Reimbursement and coverage decisions of the above cluster of DTs, come under the ambit of the **national competencies**. The structure of the respective decision-making bodies and their healthcare funders (insurance funds) remains **heterogenous**, ranging from centralized decision making that applies

³⁸⁸ Lampe K, Mäkelä M, Garrido MV, Anttila H, Autti-Rämö I, Hicks NJ, European Network for Health Technology Assessment (EUnetHTA). The HTA core model: a novel method for producing and reporting health technology assessments. *Int J Technol Assess Healthcare* 2009 Dec;25(Suppl 2):9-20

³⁸⁹ F.B. Kristensen, K. Lampe, C. Wild, M. Cerbo, W. Goettsch, L. Becla; The HTA Core Model 10 Years of Developing an International Framework to Share Multidimensional Value Assessment; *Value in Health*, 20 (2) (2017), pp. 244-250

³⁹⁰ H. Alami, P. Lehoux, Y. Auclair, M. de Guise, M.P. Gagnon, J. Shaw, *et al.*; Artificial intelligence and health technology assessment: anticipating a new level of complexity; *J Med Internet Res*, 22 (7) (2020)








³⁹¹ K. Kolasa, G. Kozinski; How to Value Digital Health Interventions? A Systematic Literature Review; *International Journal of Environmental Research and Public Health* [Internet], 17 (6)(2020)

³⁹² Bart-Jan Boverhof, W. Ken Redekop, Jacob J. Visser, Carin A. Uyl-de Groot, Maureen P.M.H. Rutten-van Mölken. Broadening the HTA of medical AI: A review of the literature to inform a tailored approach; *Health Policy and Technology*, Volume 13, Issue 2, June 2024, 100868

throughout a member state (*e.g.*, Germany, Netherlands, Belgium) or those with decentralized or regional models that start with centralized recommendations, which are then tailored to manage with regional budgets (*e.g.*, Spain, Italy). In addition, the evidence requirements to make reimbursement decisions, are likely to vary amongst national and regional actors, which is bound to delay the patients access to the benefits of DTs. A DMD task force was initiated under the French presidency of the EU, aiming to propose a harmonised evaluation framework³⁹³.

Preliminary evidence from the EDITH mapping of payers and HTA landscape, shows a rapid emergence of **fragmentation** within and between member states on the criteria, standards and the required evidence to assess digital health technologies. The governance, pathway and authority are heterogenous and remain distributed across many actors and agencies. Such fragmentations in the payer and reimbursement landscape lead to differences in the national standards and guidance. It is evident that they impact the reimbursement pathway for DTs, just like the barriers for the reimbursement of the medicines and medical devices. In the following Table, a non-exhaustive observation of health technology assessment frameworks in the context of Digital Twins in healthcare, which is likely to fall under the ambit of current-day digital health technologies (DHT) pathway for reimbursement. The organizational structure and categorisation with respect to CE marking are briefly documented for a few representative countries. Alongside, their influence and implications to the Digital Twins in healthcare and the VHT ecosystem products / solutions, are briefly explored.

Table 4: A non-exhaustive overview of several digital health technology assessment frameworks in member states, which relate to or impact the Digital Twins in healthcare, as envisioned within the VHT initiative.

Country / Dimensions	 Germany	 France	 Spain	 England	 Netherlands	 Belgium	 Finland
DHT/DMD track impacting DT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DHT/DMD HTA track	DiGA ³⁹⁴	PECAN ³⁹⁵	Digi-HTA Adapted ³⁹⁶	DTAC ³⁹⁷	Digitale Zorg ³⁹⁸	mHealth Belgium ³⁹⁹	Digi-HTA ⁴⁰⁰
CE Marking / Risk class – class 1, class 2a, class 2b or class 3	<input checked="" type="checkbox"/> Type 1, 2a <input checked="" type="checkbox"/> Type 2b, 3	<input checked="" type="checkbox"/> Type I-II-III	No explicit exclusion, covers all CE classes	No explicit exclusion, covers all CE classes	CE requested, when relevant	CE pre-requisite	No explicit exclusions
End user – Patient only, Patient with some help with HCP, Patients in consultation with HCP, only HCP-can use	<input checked="" type="checkbox"/> Patients only <input checked="" type="checkbox"/> HCP only	<input checked="" type="checkbox"/> Patients only in consultation with HCP	<input checked="" type="checkbox"/> Patients <input checked="" type="checkbox"/> HCP	<input checked="" type="checkbox"/> Patient only use & fitness tracking	<input checked="" type="checkbox"/> Patients <input checked="" type="checkbox"/> HCP	No explicit mention noted	All users
Categorization relevant for Digital Twins in healthcare	Not for HCP centric DTs for CDSS and for prevention	Early access: all Permanent: only those that do data processing & benefit	Tier C: Population & personalised DTs to inform, drive,	Tier C: Population & personalised DTs to inform, drive,	DT including algorithms, AI - robotics, prediction, AR/VR	No restriction or guideline noted	DTs comprising mHealth, AI and robotics

³⁹³ <https://eithealth.eu/news-article/press-release-digital-medical-devices-launch-of-a-european-taskforce/>

³⁹⁴ Germany DMD decision maker - <https://diga.bfarm.de/de>

³⁹⁵ France DMD program - https://www.has-sante.fr/jcms/p_3376633/fr/dispositifs-medicaux-numeriques-la-prise-en-charge-anticipee

³⁹⁶ Spain DMD assessment framework - <https://aquas.gencat.cat/web/.content/minisite/aquas/publicacions/2023/framework-adaptation-digital-hita-redets-aquas2023.pdf>

³⁹⁷ England DMD assessment framework: <https://transform.england.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/>

³⁹⁸ Netherlands DMD program: <https://www.zn.nl/dossier/digitalisering/digitale-zorg/>

³⁹⁹ Belgium DMD program: <https://mhealthbelgium.be/>

⁴⁰⁰ Finland HTA agency: <https://oys.fi/fincchta/en/digi-hita-eng/about-digi-hita/>

			diagnose and treat	diagnose and treat			
Payer / Buyer / Decision maker	BfArM G-BA	HAS CEPS	SNS	NHS	VWS ZIN	NIHDI INAMI- RIZVI	FinCCHTA COHERE Finland
Health technology assessment body	IQWiG IQTiG	CNEDiMT S CEESP	RedETS AQuAS	NICE	ZIN nictz	KCE	FinCCHTA

24.3.2 Examples of reimbursed DTs

Three representative example of DHTs in the context of clinical decision support for remote monitoring or personal health forecast were noted. Although not strongly falling in the scope of VHT, these are more than fitness or wellness apps:

- DBLG1 system⁴⁰¹, uses a hardware system in combination with software algorithm for closed-loop real-time monitoring of insulin, as well as patient vitals and nutritional uptake. It determines and administers insulin dose and has been covered for reimbursement by French agency (HAS), for a 3 year reimbursement period;
- SkinVision⁴⁰², a self-use app to detect skin cancer, using AI-based algorithm, currently reimbursed in the Netherlands;
- Kaiku Health⁴⁰³, a remote monitoring of patient symptoms and non-urgent messages for cancer patients, considered by Finnish agency (FinCCHTA).

Additional examples of **reimbursed DT technologies** were discussed in the previous chapter on regulatory science. The examples of HeartFlow™, HeartGuide™ and inHeart™ applications have received positive reimbursement decisions, in the US, UK and Japan. The example discussed in the box below, concerns DTs for preivision paediatric heart surgery, has been reimbursed by private insurers.

The above preliminary analysis of the reimbursement pathways for DMD space, showed that the fast evolving landscape of HTA frameworks could eventually internalise the assessment of VHT solutions. Hence the VHT CoP needs to take stock of the above developments and engage with the relevant stakeholders to mutually educate, inform and co-create the assessment framework and governance mechanisms.

Box 44: Success story – reimbursement of Digital Twins for precision paediatric heart surgery

Success story: Digital Twins for complex paediatric heart surgery

Status: reimbursed

Website: <https://answers.childrenshospital.org/3d-modelling-heart-surgery/>

Boston Children's Hospital has integrated computational fluid dynamics (CFD) software into its cardiac surgery program to enhance **preoperative planning for paediatric patients with complex congenital heart disease**. This innovative approach enables precise modelling of blood flow dynamics, facilitating improved surgical decision-making and patient outcomes.

Dr. Hoganson MD, a paediatric cardiac surgeon, and Dr. Hammer, PhD, a mechanical engineer, have **built a team of over 10 specialist engineers** to create and utilize imaging data to build patient-specific 3D computer models that are utilized for preoperative planning and intraoperative guidance. The Digital Twins are shared with clinicians on a digital viewing platform accessible from on any computer screen including their use in the operating room during surgery. This program includes advanced engineering analysis for **precise patch planning** in complex intracardiac and vascular reconstruction and for personalised **planning of aortic and mitral valve repair**.

⁴⁰¹ <http://doi.org/0.1111/dom.15008>

⁴⁰² <https://www.skinvision.com/>

⁴⁰³ <https://support.kaikuhealth.com/appstores/>

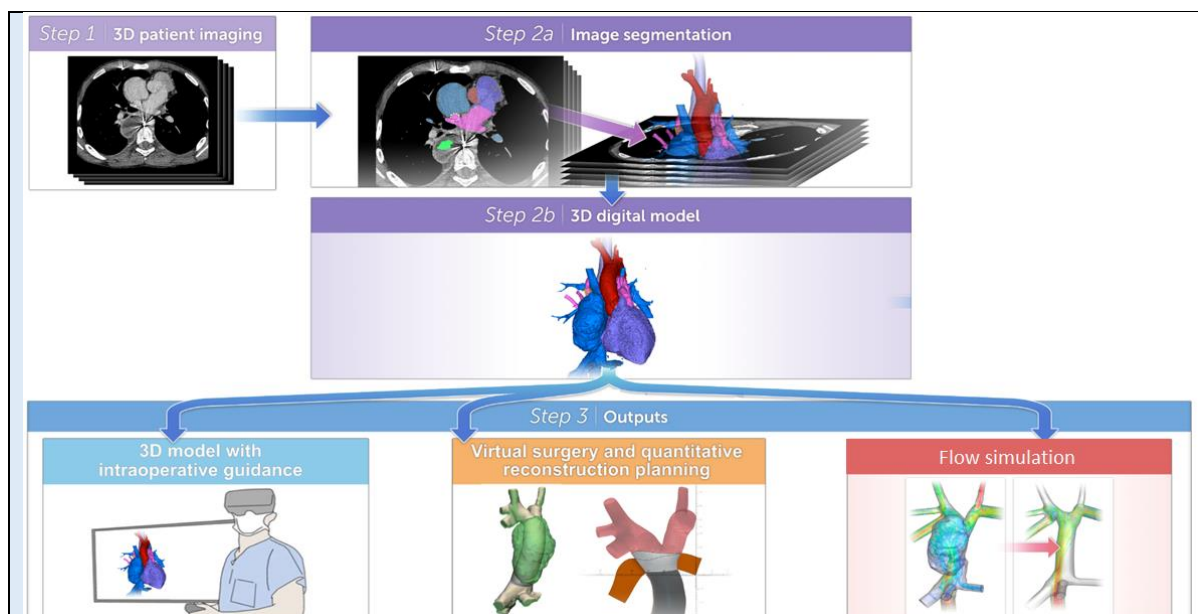


Figure 31: Boston Children's Hospital 3D modelling workflow

The team also creates Digital Twins for a small number of patients with **complex single ventricle heart disease** and in those patients flow simulation are used to plan Fontan completion or Fontan reconstruction surgery. The flow simulation helps **avoid a risk of reintervention in over 70% of Fontan patients** with high risk anatomic features which translates to a **cost savings of over \$300,000 per patient**.

In total, **more than 1,500 patients** have benefited from this patient specific 3D modelling program, the largest and most comprehensive of its kind in the world. In 2024, the team made **Digital Twins for over 35% of patients undergoing surgery at Boston Children's Hospital** with about 20% of them having virtual surgery planning and 5% having flow simulation.

Importantly, the team is now directly mentoring more than a dozen centres around the world in applying these impactful techniques. They are also spearheading **reimbursement** for these efforts, and, with the Society of Thoracic Surgeons, have applied to the Centres for Medicare and Medicaid for new CPT billing codes. They **have long been billing private health insurance** for the creation of the Digital Twins, for over 2 years.

24.4 Relevance of pre-commercial and public procurement of innovations for VHT

Similar to the conventional **supply-driven HTA** of VHT solutions presented in the previous section, there exists a complementary **demand-driven pathway for VHT technologies** to reach the end users and clinics. It is important to note that the payer and HTA community are in constant look out for technology solutions in the horizon, which meet critical unmet needs of patients and healthcare systems. Countering the flood of conventional supply-driven technology solutions, VHT ecosystem can partner with the payer/buyer community to **streamline and co-develop innovative VHT technologies**.

The healthcare innovations like those that VHT ecosystem champions are likely to fall in the horizon scanning of actors in the public procurement of healthcare innovations. For instance, initiatives to champion **pre-commercial procurement of innovations (PCP)** related to systems like telemedicine, integrated clinical decision-making systems or patient-clinician facing cascade of digital health technologies for tele rehabilitation, are already relatable to the Digital Twins in healthcare that VHT ecosystem builds upon. This could be of specific interest to the consortium of public procurement health innovations, as VHT solutions fill the current gaps within the supply-side.

On a related note, the PCPs are designed to increase the TRL levels of the DTs giving them a much-needed leverage for generating evidence that supports regulatory approval. Further, they could help VHT technologies to prepare the necessary evidence that is sought for the early market entry through **public procurement of innovations (PPI)**. Collectively these would offer the incentives to prepare for the large-scale industrialization and scale up of VHT solutions.

25 Legal aspects

25.1 Introduction

The in-depth analysis of the legal framework relevant to the VHT, which is the specific subject of the detailed and extensive assessment prepared by the EDITH-CSA⁴⁰⁴, has focused on two primary objectives:

- (i) assisting all stakeholders composing the VHT ecosystem in **identifying the key legal and regulatory challenges** that need to be overcome to ensure compliance with various applicable regulations,
- (ii) **offering specific recommendations** with the aim of helping European policymakers to better identify and remove the obstacles, detected within the complex and multi-layered legal framework, that currently hinder a wider adoption of VHT and Europe's global leadership in this field.

DT technologies and the VHT intersect with nearly all major EU data, technology, and health regulations, making them a unique test case for the coherence of European legislation in healthcare and medical research. This technology relies on vast volumes of personal and non-personal data, secondary processing of health information, AI models, and secure data exchange, while also implicating intellectual property, cybersecurity, and medical device regulations.

Given the complexity of the legal framework, codes of conduct emerge as crucial accountability tools. They bridge general legal principles with specific industry needs, reducing regulatory fragmentation and aligning with EU laws like the GDPR, AI Act, Data Governance Act, and Data Act, which encourage self-regulation to foster trust, interoperability, and compliance. **Self-imposed rules could harmonize scattered requirements** caused by varying national implementations and enhance trust among patients, users, and stakeholders. However, developing effective codes requires regulatory maturity and specialized oversight by competent authorities, which currently remains underdeveloped for VHT. To address these challenges, the analysis carried out aims to:

- (i) **help stakeholders**, especially VHT developers and users, navigate legal compliance challenges across multiple regulations
- (ii) provide recommendations for **European policymakers** to overcome barriers that slow the adoption of VHT and impact Europe's leadership in this innovative sector.

This chapter discusses 6 main areas of legislation which, to date, are relevant to the VHT initiative:

1. **Privacy and Data Protection:** GDPR Regulation (EU) 2016/679, and national associated laws;
2. **Data Governance and reuse:** the GDPR, the Data Governance Act, Regulation (EU) 2022/868, the Data Act, Regulation (EU) 2023/2854, and the European Health Data Space, Proposal 2022/0140(COD);
3. **Artificial Intelligence:** Artificial Intelligence Act, Regulation (EU) 2024/168;
4. **Medical Devices:** Medical device Regulation (EU) 2017/745 and the *In vitro* Diagnostics Medical Device Regulation (EU) 2017/746, as well as connected laws such as the Health Technology Assessment Regulation (EU) 2021/2282);
5. **Clinical Trials:** Clinical trial Regulation (EU) 2014/536 and related legislation, such as the Good Clinical Practice Directive, ICH guidelines;
6. **Intellectual Property Rights:** legislation related to copyright, database protection, trade secrets, and patentability.

Prior to the deeper dive in these 6 areas, a few cross-cutting considerations on anonymization, synthetic data and harmonization are provided.

⁴⁰⁴ <https://doi.org/10.5281/zenodo.14516807>

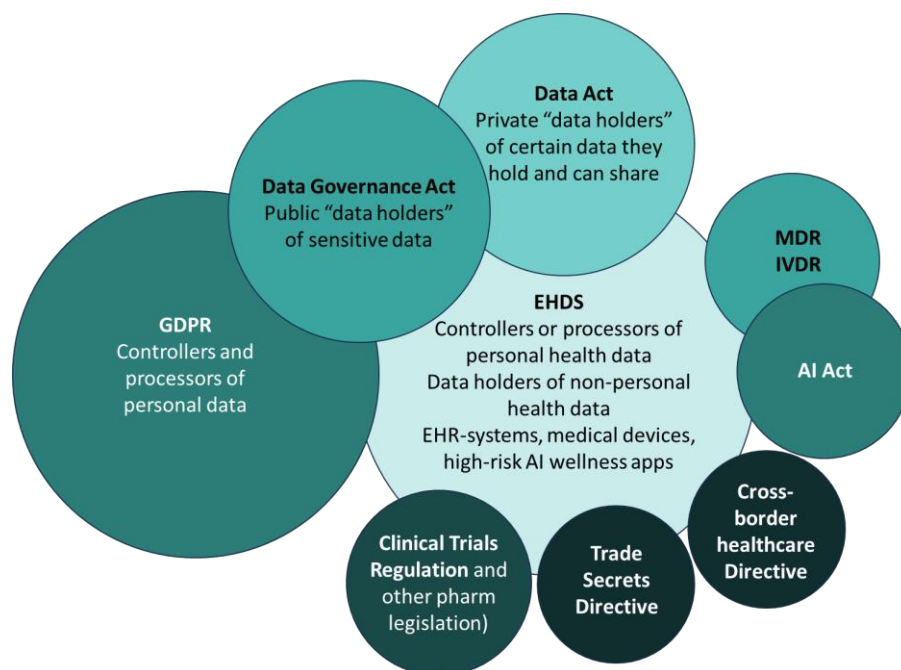


Figure 32: overview of the legislation relevant for the VHT.

25.1.1 Anonymous data and legal complexities

The legal question of when data can be considered anonymous is critical across all EU regulations, as it determines whether specific provisions apply and so directly impacts all projects involving or based on the processing of personal data. In a widely recognized article published in December 2020, Aloni Cohen and Kobbi Nissim highlighted that *“There is a significant conceptual gap between legal and mathematical thinking around data privacy”*⁴⁰⁵. This statement is especially true when it comes to data anonymization. Despite the critical importance of distinguishing between personal and non-personal data, in most cases, it remains extremely challenging to clearly differentiate between these two categories. This difficulty is anchored in both technical and legal factors. From the first perspective, the increasing availability of data points and sources, as well as the continuing sophistication of data analysis algorithms – even more in connection with machine or deep learning – and performant hardware makes it easier to link datasets and infer personal information from ostensibly non-personal data. From a legal standpoint, even after 29 years of data protection legislation (including 6 under the GDPR)⁴⁰⁶, it is **still unclear what is the crucial legal test that should be carried out to correctly qualify data as anonymous or not**.

Recital 26 of the GDPR establishes that data are **anonymous if they are not – or no longer – “reasonably likely” to be linked to an identifiable individual**. However, the Article 29 Working Party (WP29) and some national Supervisory Authorities have historically adopted a much stricter interpretation, requiring **complete irreversibility to consider data anonymous**. This rigid “zero-risk” standard has faced criticism, as it is often unrealistic in practical scenarios. For example, hospitals may retain identifiable patient data for treatment purposes while sharing anonymized data for research, making zero-risk impractical. This regulatory fragmentation creates significant challenges for organizations working across multiple Member States. **Divergent national interpretations** complicate compliance efforts, particularly in cross-border healthcare and AI-driven projects. Companies must navigate a patchwork of rules, increasing operational complexity and legal uncertainty.

The EU General Court (CJEU) addressed this issue in the *SRB v. EDPS* case⁴⁰⁷, ruling that **re-identification risk should be evaluated based on the specific recipient of pseudonymized data**, not

⁴⁰⁵ Cohen A; Nissim K., *Towards Formalizing the GDPR’s Notion of Singling Out*, 117, Proceedings of the National Academy of Sciences, 8344, 2020. <https://www.pnas.org/doi/10.1073/pnas.1914598117>

⁴⁰⁶ The Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data became applicable in December of 1995, while the GDPR entered into force in May of 2018.

⁴⁰⁷ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62020TJ0557>

other entities with (potential) access to additional information. This decision contrasts with WP29's stricter approach and aligns more closely with GDPR's "reasonable re-identifiability" standard. Indeed, the judgement offers a path forward for the VHT, by advocating a recipient-specific approach to the risk of singling-out. This shift may **enable the use of pseudonymized and synthetic data in healthcare and research** while ensuring compliance with data protection standards. Adopting advanced Privacy Enhancing Technologies (PETs) in tandem with synthetic data, especially with the addition of Differential Privacy, could significantly reduce regulatory burdens, paving the way for broader adoption of VHT. While this judgment introduces a practical framework, its final implications depend on the outcome of the appeal pending before the CJEU.

25.1.2 Synthetic Data

There are various forms of synthetic data, but the term essentially refers to the **creation of artificial data that replicates the statistical characteristics of an original dataset**. This process involves using AI-driven techniques to learn the relevant distributions of real data and then mimicking and sampling them to generate realistic, yet entirely fabricated, datasets that retain the same statistical properties as the originals. This approach enhances the protection of personal data and patient privacy while preserving the dataset's utility for statistical analysis and scientific research⁴⁰⁸.

A more methodological-oriented definition has been provided in a report specifically dedicated to synthetic data which was commissioned by The Royal Society and the Alan Turing Institute, whereby this type of data is described as "*data that has been generated using a purpose-built mathematical model or algorithm, with the aim of solving a (set of) data science task(s)*"⁴⁰⁹. Similarly, the TechSonar Report 2022-2023 issued by the European Data Protection Supervisor (a report aimed to anticipate emerging technology trends and better understand future developments in the technology sector from a data protection perspective) defines synthetic data as "*artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data*"⁴¹⁰.

Among the numerous advantages offered by this technique, the following are worth highlighting, when focusing first on population-based modelling for subsequent personalization in VHT applications:

- **De-biasing:** biases inherent in training data can be reduced by providing a unified, de-biased dataset for training models, mitigating the amplification of historical biases and supporting fairer machine learning outcomes.
- **Data augmentation and imputation:** synthetic data can expand small datasets, enhance robustness against outliers, and fill gaps in missing or skewed data distributions. It provides a cost-effective solution for generating labelled data required for machine learning, particularly in fields like scientific research, where large, high-quality datasets are essential.
- **Data minimization:** synthetic data reduces privacy risks associated with high-dimensional, sensitive datasets by creating realistic but non-identifiable datasets. This approach respects confidentiality while addressing vulnerabilities in traditional anonymization techniques.

However, **synthetic data presents its own challenges**. The quality of synthetic data depends on the accuracy of the original dataset and the algorithms used. It may also fail to capture outliers, which are often critical in clinical research. Regulators emphasize that synthetic data must balance utility and privacy, meeting high standards to ensure effective protection.

The EU AI Act has introduced synthetic data as a key regulatory element. Articles 10 and 59 treat synthetic data as equivalent to anonymized data in specific contexts, such as training and testing high-risk AI systems. This recognition reinforces the role of synthetic data as a privacy-preserving alternative to real data. However, the European Data Protection Supervisor advises caution, requiring careful risk

⁴⁰⁸ The Synthetic Data 'Industry Connections Activity Initiation Document' adopted by the IEEE Standard Association (Version: 2.0, dated 13 November 2023) states that synthetic data "(...) is highly realistic and statistically representative to the original data and thus suitable to serve as a drop-in replacement for it (e.g. for AI training). Yet – when generated with appropriate privacy mechanisms – synthetic data is fully anonymous and impossible to re-identify". <https://ieee-sa.imeetcentral.com/p/eAAAAAASqnGAAAAACXNJY>

⁴⁰⁹ Jordon, J., Szpruch, L., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., Cohen, S. N., and Weller, A. Synthetic Data – What, why and how? arXiv:2205.03257, 2022. <https://arxiv.org/abs/2205.03257>

⁴¹⁰ The TechSonar 2022-2023. https://edps.europa.eu/data-protection/our-work/publications/reports/2022-11-10-techsonar-report-2022-2023_en.

assessments to ensure compliance with GDPR standards. Factors like the generative model used, processing context, and applied safeguards must be evaluated case by case.

25.1.3 Towards a Harmonized Framework

The lack of consistent criteria for assessing data anonymity, alongside the divergent implementation of the GDPR across Member States in the health and research sector, hamper innovation in data-driven sectors. These huge hurdles are well described in the recent and comprehensive report titled ‘*The future of European competitiveness*’, published by the EU Commission, as well in several documents prepared by TEHDAS1, namely the joint action ‘Towards the European Health Data Space’, and in the ‘*Assessment of the EU Member States’ rules on health data in the light of GDPR*’ conducted by the EU Commission.

This makes the adoption of the interpretation provided in the *SRB v. EDPS* judgment even more crucial, as it introduces a flexible and practical interpretation of GDPR Recital 26, aligning legal standards with operational realities, especially in the healthcare sector. This approach encourages the safe sharing and reuse of data while maintaining robust privacy protections, fostering the whole VHT ecosystem. In parallel, synthetic data combined with PETs and effective safeguards are capable of providing a transformative solution for data privacy. By harmonizing regulatory interpretations and promoting flexible standards, the EU can facilitate innovation in AI, healthcare, and scientific research, across the VHT domain. This **balance between security and practicality** is essential for nurturing cross-border collaboration and ensuring Europe’s leadership in digital transformation and healthcare innovation.

25.2 Privacy And Data Protection

Creating Virtual Human Twin applications requires extensive data collection from various sources, such as Electronic Health Records, medical imaging, wearable devices, genomic data, behavioural information, and socioeconomic factors. Most of these datasets are highly personal and often fall into special categories under the GDPR, demanding strict safeguards for their collection, processing, and especially reuse. The VHT integrates identifiable and re-identifiable information to simulate individual-specific insights, which raises privacy and ethical challenges. **Ensuring compliance with GDPR principles** like transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity, confidentiality, and accountability is thus essential. These principles aim to protect patient privacy, maintain data security, and ensure responsible processing of sensitive information, especially given the advanced predictive capabilities of the VHT.

First and foremost, all **data processing** must be designed and carried out in compliance with the principles outlined in Article 5 of the GDPR, whose scope and significance often extend to ethical considerations.

The processing of personal data is lawful only if based on one of the legal grounds afforded by the GDPR (Art. 6), including individual consent, performance of a contract, compliance with legal obligations, protection of vital interests, and pursuing a public interest or a private company’s legitimate interests. For sensitive data like health or genetic information, Article 9 imposes stricter conditions to mitigate risks such as breaches, discrimination, or misuse. Exceptions to the prohibition of processing this kind of data include *inter alia*, in the health domain, obtaining individual explicit consent, pursuing public health interests, conduct of scientific research and provision of healthcare.

Recital 33 of the **GDPR allows broad consent for scientific research**, acknowledging that specific purposes may not be known at the time of data collection. This approach facilitates ongoing research while maintaining ethical standards, enabling controllers to adapt data use within related research areas without requiring new consent for every purpose. However, strict safeguards like pseudonymization must be in place.

Secondary use of personal data (Article 6(4)) requires compatibility with the original purpose. Factors like the relationship between purposes, the sensitivity of data, and existing safeguards determine whether the envisaged secondary processing may be considered in line with the original one. If the compatibility test fails (meaning the new purpose is not considered compatible), data controllers are required to rely on another legal basis under Article 6(1) to process the data for the new envisaged purpose, including obtaining fresh consent from data subjects, or fulfilling a legal obligation.

In this respect, **Art. 5 GDPR establishes a crucial presumption of compatibility of the reuse of data for carrying out scientific research activities, provided that appropriate data minimization measures are put in place**, such as pseudonymization or encryption of the data⁴¹¹.

Nonetheless, Article 9(4) of the GDPR almost nullifies such a flexibility, by allowing Member States to “*maintain or introduce further conditions, including limitations*” about the processing of health and genetic data. This means that while the GDPR provides a general and favourable framework and exceptions under which special categories of data can be lawfully processed within the EU research ecosystem, individual **Member States are empowered to impose stricter national regulations**, where they see fit, *de facto* undermining and limiting the conduct of medical research across the EU in several ways, including:

- **Inconsistent regulations across Member States:** Art. 9(4) leads to a lack of regulatory harmonization across the EU, making it more difficult for researchers and organizations carrying out multinational studies on the account of the need to comply with varying national laws. Indeed, researchers conducting cross-border clinical trials or medical studies may face challenges in navigating multiple sets of national regulations, which can create administrative burdens and delays.
- **Stricter conditions for health data processing:** many Member States have imposed additional limits on sharing health data. For example, extra approvals or stringent data protection measures, including sometimes obtaining authorization from competent national Supervisory Authorities, may be required before health data can be shared for international research collaborations.
- **Limiting access to data for secondary research:** some Member States have limited the reuse of previously collected health data, *e.g.* for secondary research purposes or for developing a Digital Twin in health, without obtaining new consent or explicit green light from competent Data Protection Authorities, even when the GDPR provides exceptions under certain conditions (*e.g.*, scientific research with safeguards like pseudonymization). Retrospective studies and related activities, which rely on existing health data, could be particularly affected by stricter national laws that make it difficult to access or process previously collected data.
- **Impact on genetic and biometric data research:** some Member States have introduced additional restrictions or conditions for processing genetic data, given its highly sensitive nature, thus limiting research opportunities in fields such as genomics, personalized medicine, and pharmacogenomics, where large datasets of genetic information are essential.
- **Complexities in public health research:** medical research and studies related to public health emergencies or pandemics face barriers if national laws impose stricter conditions on data collection and sharing, even when public health authorities have a strong interest in accessing this data for research purposes.
- **Potentially limiting scientific collaborations:** the introduction of varying national requirements discourages international collaboration in medical research, as institutions or researchers in Member States with stricter data protection rules might find it difficult to collaborate with researchers in countries with more flexible approaches, due to legal complexities around data sharing and compliance.

25.3 Data Governance

25.3.1 European Health Data Space

The European Health Data Space is a regulation aimed at harmonizing access to and exchange of electronic health data across EU Member States⁴¹². It establishes **rules, standards, and infrastructures for the primary and secondary use of health data to support healthcare, research, and innovation**. For VHT, the EHDS offers a robust framework that facilitates the secure integration of diverse datasets, promoting their application in personalized medicine, clinical trials, and public health research.

⁴¹¹ Helpful indications are provided, with regard to the ‘compatibility presumption’ established by Art. 5 GDPR, in the ‘A Preliminary Opinion on data protection and scientific research’ adopted by the EDPS on 6 January 2020. https://edps.europa.eu/sites/default/files/publication/20-01-06_opinion_research_en.pdf

⁴¹² Progress of the legislative process: [https://oeil.secure.europarl.europa.eu/oeil/en/procedure-file?reference=2022/0140\(COD\)](https://oeil.secure.europarl.europa.eu/oeil/en/procedure-file?reference=2022/0140(COD))

In this respect, the main benefits deriving from the EHDS may be:

- **Data integration:** easier and streamlined access to health data from various sources (*e.g.*, EHRs, clinical trials, biobanks) enhances VHT accuracy and comprehensiveness.
- **Interoperability:** standardized data formats (European Electronic Health Record Exchange format (EEHRxF)) support consistent data usage across borders, enabling scalable and refined VHT models.
- **Secondary data use:** regulated access to de-identified data accelerates innovation, allowing VHT to contribute to predictive health, treatment planning, and resource optimization.
- **Cross-border research:** leveraging diverse population data improves VHT relevance in addressing regional variations in genetics, lifestyle, and healthcare systems.

The **health data holders**, under the EHDS, are all natural or legal persons, public authorities, agencies or other bodies operating in the healthcare or care sectors (including reimbursement services), or developing products or services intended for the health or care sectors, or developing or manufacturing wellness applications, or performing research in relation to the healthcare or care sectors, or which is a Union institution, body, office or agency, whenever such natural or legal person has:

- in its **capacity as a data controller** (or joint controller) under the GDPR, the right or obligation, in accordance with applicable law, to process personal electronic health data for the provision of healthcare or care, or for public health, reimbursement, research, innovation, policy making, official statistics, patient safety or regulatory purposes; or
- the ability to make available, including to **register, provide, restrict access or exchange** non-personal electronic health data, through control of the technical design of a product and related services.

Therefore, **data holders shall act as controllers in relation to electronic health data of personal nature**, meaning that the EHDS will not apply to data holders that collect and process such data on behalf of others, in the quality of data processors. For example? hospitals, as data controllers, are data holders of their EHRs. Similarly, pharma companies are data holders of their clinical trial data and biobanks. Medical device manufacturers may be data holders of non-personal data generated by their devices, if they have access to those data and the ability to generate them, while they would not qualify as data holders in cases where they merely process personal electronic health data on behalf of a hospital. The other actors established under the EHDS are:

- **‘Data user’**, *i.e.* any natural or legal person, including EU institutions, bodies or agencies, that is granted lawful access (by a health data access body) to electronic health data for secondary use pursuant to a data permit, a data request, or an access approval by an authorized participant in HealthData@EU (the cross-border infrastructure for secondary use of electronic health data);
- **‘Health data access body’**: Member States’ national authorities that will control the use of and access to electronic health data (EHD) in their national Secure Processing Environments. Their remit is mainly to issue permits, manage access time limits and address data breaches.

One of the primary challenges the EU Legislator aims to tackle through the EHDS – and the most relevant for the VHT ecosystem – is **promoting health data reuse** by addressing the national fragmentation stemming from Article 9.4 of the GDPR. Therefore, focusing on the secondary use of EHD, it must first and foremost be pointed out that access to specific categories of data (enumerated below) must only be granted by health data access bodies to health data users where the processing by the latter is necessary for one of the following specific purposes, many of which may be highly relevant for VHT:

(i) **Public Sector and Union Bodies:**

- Public health: activities like managing cross-border health threats, public health surveillance, and ensuring healthcare safety and quality.
- Policy making and regulation: supporting public sector bodies in health policy development and regulatory tasks.
- Official statistics: generating health-related statistics at national or EU levels.

(ii) **Public and private sector entities:**

- Education and training: activities in health or care sectors, including medical and clinical education.
- Scientific research: research benefiting public health, health technology assessment, and the safety of healthcare, medicinal products, and medical devices. Examples include the development of innovative healthcare products or services, and the testing and evaluating of algorithms in AI systems, medical devices, or digital health applications.
- Improving healthcare delivery: Using EHD to enhance care delivery and treatment optimization.

Health data holders are mandated to make a vast array of data available including, *e.g.*, EHRs, genomic data, clinical trial data, public health registries, and wellness application data, aggregated healthcare data, administrative data, pathogen data, and socio-economic determinants of health. This comprehensive dataset supports VHT development in personalized medicine, predictive modelling, and public health applications.

Member States must designate one or more **Health Data Access Bodies** responsible for overseeing access to health data for secondary use. They may establish new public bodies or use existing ones, provided they meet the EHDS requirements. If multiple HDABs are designated, one must act as a coordinator within the Member State and in collaboration with HDABs across the EU. Each HDAB must contribute to the uniform application of the EHDS across the EU and cooperate with all other HDABs, as well as with the EU Commission and competent Data Protection Authorities, to ensure compliance with the GDPR. Among others, HDABs are responsible for:

- **Anonymization and pseudonymization:** providing anonymized data where feasible; pseudonymized data is permitted only when anonymization is insufficient.
- **Data permits:** issued to users who demonstrate compliance with EHDS criteria, including necessity, proportionality, and safeguards for privacy and security.
- **Cross-border cooperation:** facilitating data exchange through the HealthData@EU infrastructure, promoting interoperability across Member States.

Crucially, Recital 43 EHDS states that the EU Commission should set out the procedures and requirements, and provide technical tools, for a **unified procedure for anonymising and pseudonymising** the electronic health data under the EHDS for secondary use.

While the EHDS promotes data harmonization, national divergences allowed under GDPR Article 9(4) remain a challenge. These discrepancies affect data reuse, particularly in research contexts, creating barriers to collaboration. By addressing these regulatory inconsistencies and further clarifying anonymization and pseudonymization standards, the EHDS has the potential to significantly advance VHT development and healthcare innovation across the EU.

25.3.2 Data Governance Act And Data Act

The Data Governance Act establishes a robust legal framework for **secure and trusted data sharing** across the EU. It aims to foster innovation while safeguarding privacy, security, and intellectual property rights. As a foundational pillar of the EU data strategy, the DGA supports common European Data Spaces, including health and research, by regulating access to data held by public sector bodies, establishing frameworks for trusted data intermediation services, and encouraging data altruism initiatives for voluntary data sharing in the public interest.

When it comes to a foundational element driving the development and functionality of any VHT application, such as sharing and reusing personal health data, the DGA serves as the overarching framework, or the '*genus*', while the EHDS represents a more tailored application, or the '*species*', within this broader structure. Given that the EHDS establishes a more specific and detailed regulatory regime for health data, it takes precedence over the general provisions of the DGA.

The main actors under this Regulation are:

- **Public sector bodies:** acting as data holders with authority to grant or deny access to specific data categories under national or Union law.

- **Data intermediaries:** neutral third parties facilitating data exchanges between data holders and users while adhering to strict transparency and neutrality requirements.
- **Data altruism organizations:** entities enabling voluntary data sharing by individuals or businesses for general interest objectives like healthcare research and public policymaking.

The DGA emphasizes **neutrality, accountability, and secure environments for data sharing**, including anonymization and pseudonymization, to build trust among stakeholders. Both stand-alone organizations providing data intermediation services and multi-service companies offering intermediation alongside other activities can serve as trusted intermediaries. In the case of the latter, the data intermediation function must be legally and financially separated from all other business operations, ensuring transparency and neutrality. The types of data covered by the DGA include data protected under the following grounds: (a) commercial confidentiality, encompassing business, professional, and company secrets; (b) statistical confidentiality; (c) intellectual property rights belonging to third parties; (d) personal data, provided such data are not within the scope of the Open Data Directive. The DGA also introduces the European Data Innovation Board to **promote interoperability and consistent best practices** across Member States.

Data Act (applicable from September 2025), complements the DGA by establishing **clear rules on data access and use**, seeking to ensure fairness in the digital environment, stimulating data-driven innovation, and promoting a competitive data market. This Regulation mainly aims to (i) empower users of connected devices with access and control over the data they generate, (ii) foster business-to-business (B2B) and business-to-government (B2G) data sharing for innovation and public interest, and (iii) promote interoperability to address data lock-in challenges and enhance the EU's digital competitiveness. It promotes:

- **Access-by-design:** connected products and services must provide data access by default in secure, standardized formats.
- **Fair data sharing:** data holders must ensure transparency and cannot refuse access without demonstrating significant harm.
- **Trade secret safeguards:** confidentiality measures protect trade secrets, but misuse can result in data sharing suspension.
- **Unfair terms test:** contracts imposing unfair terms on data sharing are subject to invalidation.
- **Interoperability standards:** it mandates harmonized technical requirements to ensure seamless data use across platforms and services.

The Act **applies to raw and pre-processed data** but excludes derived insights, proprietary algorithms, or enriched datasets, maintaining intellectual property protections.

25.4 Artificial Intelligence

Artificial Intelligence is a cornerstone in the development of the VHT, enabling transformative advancements in personalized healthcare. By analysing vast datasets, AI uncovers hidden patterns and enhances predictive accuracy, allowing for (real-time) updates to certain VHT models. This supports more precise simulations of health outcomes, better-informed clinical decisions, and proactive treatment adjustments. Additionally, AI facilitates automation of routine tasks, enabling healthcare professionals to focus on complex responsibilities, while improving patient engagement by visualizing health data in ways that foster understanding and adherence to treatment plans.

The **AI capabilities integral to VHT span various domains**, from enhancing system assessments and personalizing models to developing fully data-driven predictive systems. AI-driven models not only streamline complex processes like image segmentation but also support hypothesis generation, guiding the evolution of mechanistic multilevel models. These applications underscore AI's critical role in enabling VHT to deliver actionable insights into health and illness.

The novel EU Artificial Intelligence Act introduces a regulatory framework that profoundly impacts the development and deployment of AI systems within the VHT ecosystem. This legislation emphasizes safety, ethics, and innovation, balancing the need for regulation with fostering technological growth.

By **establishing a risk-based classification**, the AI ACT tailors obligations and requirements according to the potential harm posed by AI systems.

A primary challenge under the AI ACT is determining whether VHT or its components qualify as “AI systems”⁴¹³. The AI Act’s definition highlights autonomy, adaptability, and inference capabilities, all of which align with VHT functionalities. The ability of VHT applications to learn from new data and generate actionable outputs reinforces their likely classification as a high-risk AI system, particularly when associated with medical devices regulated under EU harmonization laws. These classifications impose stringent obligations, such as *inter alia* ensuring data quality, enabling human oversight, and maintaining transparency about system capabilities and limitations.

In addition to the definition of “AI system”, the notion of “*general-purpose AI model*” might also be relevant in the context of the use of AI technologies for the implementation of VHT technologies⁴¹⁴.

Exemptions within the AI ACT, such as those for **systems developed for the sole purpose of scientific research**, provide critical leeway for innovation. These derogations allow VHT-related technologies to be explored and refined without the immediate burden of compliance. However, once such systems transition to deployment, they must meet the requirements under the AI Act.

Stakeholders in the **VHT ecosystem**, including hospitals, researchers, and technology providers, must **navigate their roles as either providers or deployers of AI systems**. Providers are responsible for ensuring the accuracy and fairness of training and validation data, while deployers must assess the impact of AI systems on fundamental rights, especially in healthcare settings. Collaborative efforts, such as consortia involving multiple entities, introduce complexities in defining responsibilities and ensuring compliance.

The AI ACT also includes provisions to **support innovation**, such as the establishment of **AI regulatory sandboxes**. These controlled environments allow stakeholders to test AI systems under real-world conditions while ensuring adherence to ethical and legal standards. Further processing of personal data for public interest purposes, including healthcare advancements, is explicitly supported within these sandboxes.

In conclusion, AI technologies are fundamental to the advancement of VHT, offering unparalleled opportunities to enhance healthcare outcomes through personalization and predictive modelling. However, their deployment within the framework of the AI ACT introduces regulatory challenges that require careful navigation. By adhering to these regulations while leveraging opportunities for innovation, stakeholders can ensure the ethical and effective integration of AI into VHT, paving the way for a more robust and human-centered healthcare future.

Part of the European AI Strategy is the proposal for a Directive on adapting non-contractual civil liability rules to AI, the so-called **AI liability Directive 2022/0303(COD)**⁴¹⁵. The Directive aims to set general conditions for the imputability of damages arising from AI. Over the years, there have been many proposals regarding damages. Strict liability, concerning the developers of artificial intelligence, has been envisaged on the basis of liability arising from dangerous activities. However, the proposal settles on the **imputability of the damage to the provider or user on the basis of risk management**. In this respect, it establishes the conditions under which the causal link between the fault of the defendant in a damages action and the output or failure to output produced by the AI system is presumed to exist. Specifically, the claimant has to demonstrate specific elements provided by the Proposal, *e.g.*, the non-compliance with a duty of care provided by the AI Act from the defendant and causal link between the output produced by the AI system and the damage. The liability regime outlined in the AI Liability Directive clarifies and completes the framework of entities responsible for the manufacture, distribution, and use of AI. It therefore represents a useful and fundamental tool for the purposes of the

⁴¹³ «AI system’ means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments» (Article 3(1)(1) – AI Act).

⁴¹⁴ «General-purpose AI model’ means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market» (Article 3(1)(63) – AI Act).

⁴¹⁵ [https://oeil.secure.europarl.europa.eu/oeil/en/procedure-file?reference=2022/0303\(COD\)](https://oeil.secure.europarl.europa.eu/oeil/en/procedure-file?reference=2022/0303(COD))

development and distribution of AI-devices for medical purposes, VHT included. In a recent report⁴¹⁶ from the European Parliamentary Research Service (EPRS), it was proposed to extend the scope of this directive to include general-purpose and other 'high-impact AI systems', as well as software and introducing a mixed liability framework that balances fault-based and strict liability. Finally, the **study recommended transitioning from an AI-focused directive to a software liability regulation**, to prevent market fragmentation and enhance clarity across the EU – which is highly relevant for the non-AI-based DTs.

25.5 Medical Devices

The EU Regulations 2017/745 (MDR) and 2017/746 (IVDR) establish **stringent legal frameworks governing medical devices (MD) and *in vitro* diagnostic devices (IVD)** within the Union, ensuring safety, quality, efficacy, and transparency. These regulations explicitly include software, such as stand-alone applications, under their scope if certain criteria are met. This inclusion has significant implications, given that **VHT developments may be classified as Software as a Medical Device (SaMD)** under the MDR if they independently process data to generate outputs intended for medical purposes and influence decisions related to diagnosis or therapy (*cfr.* regulatory chapter). Similarly, VHT applications can fall under the IVDR when designed for examining human-derived specimens to provide diagnostic or monitoring information. Both frameworks apply rigorous pre- and post-market requirements.

The **MDR aims to ensure device quality, safety, and transparency**, supported by a clear definition of medical devices, comprehensive compliance requirements for manufacturers, and a risk-based device classification into four categories determining regulatory controls, with higher-risk devices requiring strict conformity assessments by independent notified bodies. Notified bodies are designated to assess compliance with MDR requirements and manufacturers must continuously monitor devices in real-world use, supported by a Unique Device Identification (UDI) system for traceability.

The dynamic nature of some VHT applications, characterized by adaptive algorithms and real-time updates, conflicts with MDR's static pre-market evaluation processes. Current post-market surveillance frameworks are similarly inadequate for evolving technologies like those integrating Digital Twins. Moreover, VHT's broad predictive modelling purposes often **do not fit neatly within the medical device definitions under MDR, creating regulatory uncertainty**.

The IVDR parallels the MDR in its **rigorous requirements for IVD**, with risk-based classification guiding conformity assessments. Higher-risk categories involve more stringent evaluations, including oversight by notified bodies and performance studies to ensure analytical and clinical validity.

If a VHT application integrates AI systems, the AI Act applies additional considerations, classifying such devices as "high-risk" and subjects it to notified body assessments (*e.g.*, Classes IIa, IIb, III under MDR or B, C, D under IVDR). This classification imposes further obligations on manufacturers to ensure data quality, human oversight, and ethical compliance. The adaptive and evolving nature of some VHT applications raises **challenges in meeting traditional clinical validation standards**, as their probabilistic outputs and machine learning processes deviate from conventional metrics of efficacy and safety.

Clinical validation challenges, dynamic adaptability, ambiguity in risk classification, and post-market surveillance, are all barriers complicating VHT's alignment with MDR and IVDR.

The inherent complexities of VHT applications, particularly their reliance on AI and dynamic modelling, underscore the necessity for tailored regulatory frameworks. The **MDR and IVDR, while comprehensive, must evolve to accommodate the unique attributes of technologies like Digital Twins**, ensuring that they can be effectively assessed and integrated into the EU medical device market without compromising safety or innovation.

25.6 Clinical Trials

Digital Twins represent a transformative innovation in clinical trials, reshaping traditional methodologies by simulating patient models that closely mimic real individuals or populations. These

⁴¹⁶ [https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU\(2024\)762861](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2024)762861)

advanced simulations integrate predictive analytics and data from sources like EHRs, wearables, and biosensors to predict treatment outcomes, identify risks, and tailor interventions. By enabling digital testing, VHT enhances precision and efficiency, reduces costs, and shortens trial durations, driving a shift toward precision medicine. Additionally, VHT applications' capacity for adaptive trial designs allows researchers to adjust methodologies in real-time based on continuous data analysis, optimizing patient care and trial outcomes.

The EU Clinical Trials Regulation 536/2014 **indirectly applies to VHT applications when used to inform trial design, predict outcomes, or support regulatory submissions**. Although the CTR does not explicitly address Digital Twins, its provisions become relevant when these technologies are employed alongside traditional interventional trials. In such cases, compliance with ethical standards, data integrity, patient safety, and scientific validation is mandatory. However, the CTR **limits the role** of fully AI-driven methodologies, positioning Digital Twins as **complementary tools** rather than replacements for human trials. The overall VHT initiative contributes to clinical trials in several key areas, such as:

- (i) **direct involvement in trials** by simulating patient responses or optimizing trial protocols;
- (ii) **supplementary evidence**, supporting regulatory submissions by enhancing data comprehensiveness;
- (iii) **adaptive trial designs** that enable real-time adjustments through continuous simulation and analysis;
- (iv) **post-marketing surveillance** in Phase IV trials, predicting long-term risks and monitoring drug performance;
- (v) **regulatory and operational challenges**.

The CTR, while promoting harmonized and streamlined clinical trials across the EU, presents several **challenges for the integration of VHT technologies**, including regulatory uncertainty, acceptance of virtual evidence, validation standards, data integration issues, ethical and consent concerns, as well as harmonization across Member States.

The European Medicines Agency has acknowledged the potential of digital methodologies, issuing guidelines to facilitate innovation in drug development and trial designs. These include recommendations on physiologically-based pharmacokinetic modelling and qualification of novel methodologies. However, these guidelines remain voluntary and insufficient for establishing comprehensive standards for Digital Twins.

25.7 IPR Management

The **management of intellectual property rights within the VHT ecosystem** represents a multifaceted challenge, requiring the alignment of diverse stakeholders' interests while fostering innovation, transparency, and long-term sustainability. At the core of this effort is the **need to reconcile the contributions and expectations** of researchers, businesses, healthcare providers, and patients **with the collaborative and open nature** of the ecosystem. This task involves not only addressing "background IP", which encompasses pre-existing data, models, and tools contributed by stakeholders, but also "foreground IP", which includes new assets developed through collaborative processes within the ecosystem.

One major challenge lies in **identifying which elements** of health data and related infrastructure **can be protected** by intellectual property rights and the appropriate mechanisms to do so. Individual health data, while critical, typically lack the creative input necessary for copyright protection and are considered "real-world facts". However, structured databases and synthetic data offer potential avenues for IPR. Synthetic data may qualify for copyright if sufficient creative input is demonstrated. This potential hinges on resolving authorship questions, as the involvement of AI in generating synthetic data complicates traditional notions of creativity and ownership. Similarly, databases can be protected under copyright for the creative selection and arrangement of their content, and they also benefit from *sui generis* protection under EU law for the significant investments required in their creation.

Trade secrets offer another layer of protection, particularly for databases that hold significant commercial value. Nonetheless, individual health data rarely meet the secrecy or value requirements for trade secrets. This distinction is important as the European Health Data Space introduces a mandate for sharing certain electronic health data for secondary use, even if they are protected by IPR or trade

secrets. While this promotes access to diverse datasets for research and innovation, it also introduces risks for stakeholders, who may hesitate to invest in new technologies or data-driven solutions if proprietary information must then be disclosed. Furthermore, the broad definitions and requirements of the EHDS may create legal uncertainties regarding the scope of shared data and the extent of IPR protection, potentially leading to inconsistent interpretations across Member States.

Software and AI tools form another critical aspect of the VHT ecosystem, particularly predictive models and computational tools used in simulations. These tools may be protected under copyright as literary works or patented if they address technical problems with novelty, inventive steps, and industrial applicability. However, the “black box” nature of many AI systems, where their internal workings are opaque even to developers, presents challenges for meeting the sufficiency of disclosure required for patents. Applicants must carefully navigate the line between providing enough detail to satisfy legal requirements and safeguarding proprietary knowledge. This tension is especially acute when AI models rely on training data and methodologies that are central to the invention but cannot be easily disclosed without compromising the competitive advantage.

To address these complexities, the EHDS would benefit from the establishment of **clear, uniform guidelines at the EU level** regarding data sharing and IPR protection. **Transparency** in the processes employed by Health Data Access Bodies to determine the scope of protected data is critical. Licensing agreements that require commercial entities to reinvest in public health or share benefits with public institutions could help align commercial use with ethical and societal goals.

Open-source models offer a promising alternative by fostering collaboration and reducing barriers to innovation, as the underlying principles, which grant users access to modify and redistribute software or datasets, align well with the goals of transparency and accessibility. However, implementing open-source models for AI systems is inherently more complex than for traditional software due to the interdependent components involved, such as trained models, data, preprocessing scripts, and algorithms. Clear definitions and guidance from institutions, as exemplified by the AI Act, are essential for clarifying the scope of open-source licenses and encouraging their adoption. Incentives such as compliance exemptions or reduced oversight requirements for open-source systems could further promote this model, though careful consideration must be given to downstream implications, including the limitations these licenses may impose on commercialization.

Patent protection remains a valuable tool for safeguarding innovation within the VHT ecosystem. However, meeting sufficiency of disclosure requirements for patents related to AI technologies necessitates innovative approaches. To address this hurdle, the European Patent Office could adopt **supplementary guidelines or establish secure repositories** for storing sensitive materials such as training datasets and algorithms, in analogy to practices used for biological patents. This would enable inventors to satisfy legal standards without compromising proprietary information. Moreover, international collaboration, particularly through the World Intellectual Property Organization, could harmonize patent standards, ensuring consistent global practices.

The **dual role of patents as both an incentive for innovation and a potential barrier to accessibility** underscores the importance of balancing exclusivity with collaboration. Implementing a framework inspired by Standard Essential Patents could facilitate equitable licensing while ensuring broad access to critical technologies. Such a model would be particularly relevant for foundational AI tools or methods essential to the scalability and interoperability of VHT systems. Licensing agreements under fair, reasonable, and non-discriminatory (‘FRAND’) terms could prevent monopolistic practices while encouraging innovation and collaboration.

Ultimately, an effective IPR framework for the VHT ecosystem must balance the need for openness, collaboration, and innovation with robust protections for proprietary interests. This balance requires **clear policies, transparent processes, and equitable licensing models** that foster trust, align with ethical standards, and support the long-term sustainability of the ecosystem. By addressing these challenges comprehensively, the VHT framework can maximize its potential to advance healthcare while maintaining societal trust and promoting public health objectives.

25.8 Digital Twin use cases

Box 45: Use Case - Glycemic control in ICU patients (legal)

Use Case: *Glycemic control in ICU patients*

Status: *start-up company, solution in clinical trial*

Website: www.insilicare.com

InSiliCare's AI-powered Digital Twin is a clinical decision support solution for ICU patients requiring glycemic control. It provides personalized, safe, and effective management of blood glucose levels and nutrition delivery. Insulin and nutrition treatments are calculated to maximise safety from hypoglycemia, while controlling patient blood glucose levels and optimizing nutrition towards a configurable physician-determined practice of care.

The full AI-based clinical decision support solution must be medical device regulation EU 2017/745 compliant prior to commercial use. More specifically, it likely qualifies as a class IIb active (software) medical device. As such, it will also comply with the following non-exhaustive list of identified standards and regulations: ISO13485 (quality management system), ISO14971 (risk management), IEC62304 (software lifecycle), IEC62366 (software usability), GDPR (sensitive data protection), and the AI-act (risk management of AI). This full solution methodology is patented (US11135368B2).

Box 46: Use Case - Osteoporotic fracture risk prediction (legal)

Use Case : *Bologna Biomechanical Computed Tomography for osteoporotic fracture risk prediction*

Status: *clinical research studies*

Bologna Biomechanical Computed Tomography (BBCT) is a Digital Twin methodology designed to predict the mechanical strength of the femur under critical loading conditions in osteoporotic patients. Quantitative Computed Tomography (QCT) scans of the hip region and patient data inform a subject-specific Finite Element (FE) model able to predict the risk of hip fracture at the time the CT is performed (ARF0).

BBCT processes personal data of patients in Italy and is subject to GDPR (General Data Protection Regulation). It involves special categories of personal data, such as health data, necessitating a legitimate basis for processing under Articles 6 and 9 of GDPR. The current approach relies on consent, which must meet specific requirements or compliance with Italian privacy laws. Additional compliance measures include a privacy policy (Art. 13 GDPR) and a Data Protection Impact Assessment (DPIA, Art. 35 GDPR) to evaluate risks to data subjects. Adherence to privacy-by-design and privacy-by-default principles (Art. 25 GDPR) and the use of pseudonymisation are also essential. BBCT, potentially classifiable as a high-risk AI system under the AI Act, is currently used for scientific research and not marketed, so the AI Act does not yet apply. Future development might utilise AI regulatory sandboxes (Art. 57 AI Act). BBCT could benefit from "data altruism" under the Data Governance Act and, potentially, from the European Health Data Space (EHDS), which would enable secondary use of health data under GDPR. However, the EHDS has not yet been finalised, and its impact remains to be assessed. If licensed, BBCT must first be certified under the Medical Device Regulation (Reg. 2017/745) and liability standards (Dir. 85/374/EC). Compliance with GDPR for data transfers and adherence to future AI Act and AI Liability Directive rules will also be required.

Box 47: Use Case - Epileptogenic zone localisation for surgical planning in epilepsy patients (legal)

Use Case : *Epileptogenic zone localisation for surgical planning in epilepsy patients*

Status: *in clinical use*

Website: <https://www.cloudsofcare.com/>

Persyst ESI powered by Epilog is a neuroimaging solution that automatically combines scalp EEG data with a patient's MRI to perform Electrical Source Imaging (ESI). It pinpoints the origin of brain activity linked to seizures, helping clinicians accurately localize the epileptogenic zone—critical for surgical planning in epilepsy patients.

Access to Epilog's ESI processing pipeline is integrated in the Persyst software platform, allowing users to upload EEG and MRI data directly to the Persyst ESI powered by Epilog secure cloud server. EEG and MRI data are selected via the local Persyst ESI application and are automatically deidentified prior to upload to the Persyst ESI powered by Epilog secure cloud server. No protected health information leaves the local network. After the ESI results are ready, the user downloads the deidentified results to the local network, where patient

identifying information is added back via the local Persyst ESI application. The final reports are then ready for upload to the hospital's electronic health record system.

Persyst ESI powered by Epilog is owned by Clouds of Care NV, an ISO 13485:2016 and ISO 27001:2013-certified MedTech company in Gent, Belgium housing CE-marked and FDA-cleared electrophysiology applications for clinical care, and tailor-made solutions for clinical trials and research. Persyst, a MedTech company in San Diego, USA, has the exclusive marketing rights in Europe and USA. Clouds of Care NV is a spin-off company of Ghent University, Imec and Antwerp University in Belgium. Epilog is a trademarked brand of the Clouds of Care digital health technology portfolio reflecting the presence in the epilepsy market segment.

Box 48: Use Case - the Atrial Modelling Toolkit for cardiovascular Digital Twins (legal)

Use case : the Atrial Modelling Toolkit for cardiovascular Digital Twins

Status: for research purposes

Website: <https://github.com/pcmlab/atrialmtk>

The Atrial Modelling Toolkit⁴¹⁷ (**atrialmtk**) aims to overcome the challenges of constructing cardiac models at scale through the development of a robust, open-source pipeline for bilayer and volumetric meshes for atrial models.

GDPR: Researchers developing the atrialmtk tools and applications follow data management protocol, standards and guidelines set out by the Health Research Authority, Barts Clinical Trials unit and the Cardiovascular Devices Hub. Models are based on processing of anonymised patient data. Patients have provided consent to the clinical team for using their data for research purposes. Patient data is anonymised in hospital before transfer to the modelling team/upload to data source with the hospital team (NHS Barts⁴¹⁸) maintaining a linked list (pseudo-anonymised data). The data is then downloaded and stored with data transfer agreement (where applicable/sent directly from the clinical team) and an expiry date by which data must be deleted. Data management workflow documents are in place for data governance to detail data access, data anonymisation workflows, decision to release process and the national data opt-out⁴¹⁹. The developed models are released in accordance with protocols set out by the Health Research Authority and local organisation. Rare cases will not be published, a prevalence rate of 10 per 10,000 will be used as a cut-off value for identifiability.

25.9 VHT-related remarks and recommendations based on current EU policies

Policymakers must carefully address several significant regulatory barriers embedded within the above-mentioned EU legal frameworks to **facilitate the broader adoption and integration of DTs** into the healthcare and research sector. These barriers, which range from inconsistencies in definitions and overlapping requirements to the lack of clear guidelines for emerging technologies, present a complex landscape that demands a coordinated and comprehensive policy response.

To achieve this, it is essential to **harmonize key aspects of regulations** such as the GDPR, EHDS, AI Act, Data Act, and MDR/IVDR. Policymakers should prioritize resolving ambiguities in the classification and application of these frameworks, particularly in terms of risk assessment, data governance, and compliance pathways. Clear and consistent guidelines are needed to ensure that VHT solutions can be effectively validated, regulated, and integrated into healthcare systems without being stifled by duplicative or outdated regulatory requirements.

Additionally, **fostering innovation through targeted measures** such as regulatory sandboxes, interoperable data standards, and support for Privacy-Enhancing Technologies (PETs) can help bridge the gap between current regulations and the dynamic capabilities of VHT. In this context, the role of policymakers is not only to address these barriers but also to create an enabling environment that supports the responsible development and deployment of VHT as a cornerstone of digital health innovation in Europe.

⁴¹⁷ <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2023.0038>

⁴¹⁸ <https://www.bartshealth.nhs.uk/>

⁴¹⁹ https://www.cemrg.co.uk/models_files/Cardiac%20Modelling%20Data%20Management%20Workflow%20and%20Data%20Types_V1.0.0_21-12-21.pdf

The EDITH legal landscape report⁴²⁰ provides an extensive list of challenges and recommendations. Among the most pressing and significant challenges that need to be addressed, the following deserve particular attention.

A. General Data Protection Regulation

The GDPR determines several hurdles due to divergent interpretations by Member States regarding data anonymization. Current uncertainties hinder innovation in sectors like health and research, particularly concerning the VHT ecosystem. The high threshold for anonymization, based on absolute irreversibility, limits data usability, increasing compliance costs and legal risks.

Proposed Solutions:

- a) Develop EU-wide standardized guidelines on anonymization and pseudonymization, incorporating globally recognized technical standards (*e.g.*, ISO/IEC).
- b) Shift from an "absolute anonymity" approach to a more practical "relative anonymity" model, where re-identification risks are contextualized.
- c) Promote Privacy-Enhancing Technologies, such as Secure Multi-Party Computation and Homomorphic Encryption to bolster data minimization and security, and Synthetic Data and Differential Privacy to foster anonymous data generation.
- d) Address the fragmented use of "broad consent" for scientific research by providing clearer regulatory guidance and promoting ethical safeguards, ensuring trust and transparency.

B. European Health Data Space (EHDS)

Despite the EHDS's intention to harmonize the secondary use of health data, Member States retain significant discretion under GDPR Article 9.4 to impose additional restrictions, creating a fragmented landscape that can still undermine data reuse objectives.

Proposed Solutions:

- a) Strengthen enforcement mechanisms to prevent national limitations on secondary health data use.
- b) Establish detailed guidelines for Health Data Access Bodies on applying anonymization and pseudonymization techniques, addressing existing uncertainties.
- c) Clearly define permissible research-related purposes for secondary use, ensuring clarity on the scope of activities and products that qualify.

C. Artificial Intelligence Act

The AI Act adopts a broad definition of "AI system" leading to potential uncertainties and inconsistencies, especially in the context of VHT. The lack of clear boundaries for the "scientific research exception" further complicates applications.

Additionally, the potential overlap between the risk classification established by the AI Act and that of the Medical Device Regulation raises concerns about duplicative compliance requirements and regulatory inefficiencies, further emphasizing the need for harmonized guidelines.

Proposed Solutions:

- a) Develop comprehensive guidelines to clarify the definition of AI systems, using practical examples, particularly for the VHT.
- b) Define the scope of "sole purpose of scientific research and development" for AI exemptions.
- c) Harmonize risk classification rules between the AI Act and related regulations, such as the MDR and IVDR, to streamline compliance and avoid duplication.
- d) Issue guidelines to facilitate compliance with high-risk system requirements, tailored to strategic domains like healthcare.
- e) Foster regulatory sandboxes and encourage the development of sector-specific codes of conduct for VHT.

⁴²⁰ <https://doi.org/10.5281/zenodo.14516807>

D. Data Act and Data Governance Act

The definitions of “data holder” across the Data Act, DGA and EHDS are inconsistent, creating operational and legal ambiguities, particularly for entities subject to multiple regulations.

Proposed Solution

- a) Harmonize these definitions or provide detailed guidance to ensure consistency and clarity in their application across regulatory contexts.

E. Clinical Trials Regulation

Digital Twins and other *in silico* technologies lack clear regulatory pathways under the CTR. Unlike traditional trials, there is no standardized approach for validating these technologies or incorporating them into trial processes.

Proposed Solutions:

- a) Develop EU-wide guidelines on validation and regulatory acceptance for Digital Twins in clinical trials, emphasizing reliability and reproducibility.
- b) Establish a standardized European pathway for iterative and adaptive models, enabling dynamic updates without full re-certification.
- c) Encourage interoperability and data integration, leveraging frameworks like the European Electronic Health Record Exchange Format (EEHRxF).

F. Medical Devices Regulations (MDR/IVDR)

Conflicting risk classifications under the AI Act and MDR/IVDR create duplication of compliance requirements for AI-enabled medical devices like DTs. Moreover, the current regulatory framework for medical devices is unsuited for VHT’s dynamic and evolving nature, complicating safety and performance assessments.

Proposed Solutions:

1. Harmonize risk classification and streamline requirements through joint guidelines by regulatory bodies to prevent redundancy and ease compliance.
2. Adapt regulatory pathways to accommodate the dynamic nature of VHT, ensuring safety and compliance while promoting innovation.

25.10 Towards a VHT code of conduct

The examination of the challenges arising from EU legislation governing VHT and their associated stakeholder ecosystem highlights the need for the **creation of a comprehensive Code of Conduct** pursuant to Article 40 of the GDPR. Such a Code, complemented by essential policy initiatives in areas like data reuse, health systems interoperability, and AI-driven technologies, is critical for fostering informed, cohesive, and widespread adoption of this transformative technology. This vision aligns with broader goals of ensuring regulatory clarity, ethical governance, and technological innovation while addressing pressing societal challenges in healthcare.

A Code of Conduct serves as a **pivotal accountability tool tailored to the unique needs of the VHT technologies**, offering clarity and harmonization across the EU. It bridges the gap between general legal provisions and sector-specific realities, promoting consistency, trust, and compliance. Such a framework would outline best practices and acceptable standards for data use, enhancing transparency and accountability for stakeholders and data subjects alike. From a procedural perspective, the drafting and the implementation of a sector-specific Code involve collaboration among associations, data controllers and processors. These stakeholders would propose rules addressing fair processing, transparency, anonymization, data breaches, cross-border transfers, and dispute resolution. Following supervisory authority approval and a European Data Protection Board (EDPB) review, the Code could attain EU-wide binding status, unifying data governance and protection standards across Member States.

While a comprehensive Code of Conduct offers a long-term solution, **immediate regulatory actions are essential to address pressing challenges**. The fragmented regulatory landscape imposes significant burdens on technology and data-driven markets, deterring investment and undermining

competitiveness. Reducing overlapping regulations, streamlining compliance pathways, and fostering a cohesive regulatory environment are critical for mitigating these barriers.

The EHDS provides a unique opportunity to **harmonize guidelines and procedures** across EU institutions. Establishing a strategic committee within the EU Commission would enable coordinated oversight and policy development, ensuring alignment across directorates, executive agencies, and stakeholders. This approach would address interdependencies between regulations, such as GDPR, the AI Act, MDR/IVDR, and CTR, while fostering innovation in research, healthcare, and technology.

The EHDS is positioned to enable the secondary use of health data, a cornerstone for developing the VHT initiative. By facilitating access to high-quality, interoperable, and ethically managed health data, the EHDS can unlock transformative benefits in predictive modelling, treatment optimization, and patient care. Targeted policy adjustments and EU-wide best practices would create a robust data ecosystem that supports the VHT initiative while ensuring compliance with GDPR and other regulatory frameworks. The strategic committee envisioned under the EHDS would engage with national authorities, stakeholders, and advisory groups to address barriers identified in prior analyses. By promoting interoperability, technical standards, and ethical principles, the committee would enhance regulatory clarity and foster innovation.

Effective intellectual property rights management is essential for balancing data-sharing obligations with the need to incentivize private investment in the VHT ecosystem. Copyright, database protection, trade secrets, and patentability must be harmonized to safeguard innovation while promoting accessibility. Open-source approaches could enhance collaboration but require careful regulation to protect sensitive data. Challenges such as the “black box” nature of AI complicate patentability under European Patent Convention requirements. Creating secure repositories for AI models, training data, and methodologies would address disclosure issues. Licensing frameworks inspired by Standard Essential Patents, operating under Fair, Reasonable, and Non-Discriminatory terms, would ensure broad accessibility while compensating innovators. A harmonized IPR policy for VHT would prioritize transparency, equitable licensing, and alignment with public health objectives, fostering collaboration and sustainable innovation.

The **EU must seize the opportunity to establish itself as a global leader in the VHT domain**, particularly as obligations under the AI Act and EHDS take effect around 2028–2030. By addressing regulatory fragmentation, fostering trust, and encouraging investment, the EU can position the VHT as a cornerstone of healthcare innovation. Policy-wise, the VHT represents a “growth-enhancing expenditure” within the EU’s economic framework, supporting sustainable healthcare systems and welfare models. By **integrating these technologies with a cohesive regulatory environment**, the EU can unlock their full potential, driving research innovation, economic growth, and global competitiveness.

In light of the above, a dual-track approach is essential for supporting the responsible adoption of VHT. Immediate regulatory actions must address current barriers to innovation and investment, while the gradual development of a comprehensive Code of Conduct will ensure long-term regulatory clarity and ethical integrity. By fostering a unified and accessible data ecosystem under the EHDS, the EU can nurture a robust VHT landscape that balances innovation with societal values, positioning itself as a leader in the global AI-driven healthcare market.

26 The ethical dimension of the VHT

Individuals' decision-making and the functioning of technology guided by personal values can be influenced by **ethical principles**, which serve as a framework of moral standards. Ethical obligations can provide support and motivation to stakeholders, ensuring the safeguarding and advancement of human values. While legal frameworks establish the formal prerequisites for individuals and organisations, ethical obligations can act as catalysts for what is stipulated by the law. Moral considerations aid in shaping people's actions, imposing limits, and providing guidance for technological advancements. Moreover, ethics can form the foundation of laws, with legislation often originating from ethical quandaries. For instance, concepts like informed consent, privacy, and confidentiality may be intertwined with the principle of individual autonomy. Similarly, anti-discrimination laws can be regarded as stemming from the principle of justice. However, no matter how much the law endeavours to uphold ethical concerns, it may not always prevent morally undesirable consequences from occurring.

Certain **technologies** have **sparked ethical debates** and **necessitated significant policy initiatives** aimed at controlling their trajectory, typically only after some harm has been caused. By identifying and examining potential concerns, ethics can serve as a tool to assess the risks or advantages associated with tangible forms of harm that may arise from emerging technologies. This is particularly relevant in the case of Artificial Intelligence (AI) and its regulation. Ethical examinations of AI have highlighted potential issues that may require regulation, such as bias and opacity within AI systems. To illustrate, this prompted the EU legislator to outline specific requirements regarding AI in the AI Act. The Highest authorities in the United Nations (**IFAP Council** at UNESCO) are advocating the need to accompany technological development with political, legal and ethical guidelines. Among the main ethical principles often cited are human control, transparency, responsibility in the use of technology, and inclusion. UNESCO also stresses that technological literacy is crucial and includes education and accessibility, which are key elements in this technopolicy development. Global governance is needed, and it is important that close, cross-disciplinary, cross-cultural collaborations are envisaged.

While offering various benefits such as improved diagnostics and clinical precision, less invasive treatments, faster drug discovery, and optimized regulatory processes, **VHT also introduces significant ethical challenges**. These revolve around privacy, patient autonomy, data accuracy, fairness, technological over-dependence, and ownership, compounded by the integration of multiple emerging technologies. The **ethical framework for VHT** must not only complement legal requirements but also **proactively address societal concerns** through transparency, inclusiveness, equality and accountability. Early ethical scrutiny is critical to guiding the responsible development of this technology. Some of the fundamental principles that warrant careful consideration include a range of values and guidelines designed to ensure the responsible development and deployment of this technology. Overall a multidisciplinary approach is needed to address ethical challenges in *in silico* medicine.

26.1 Privacy, accuracy, and ownership in VHT ethics

26.1.1 Privacy and data protection

Privacy emerges as a critical ethical concern due to the vast amount of sensitive data that VHT requires. Aggregating clinical and non-clinical data allows for deeply personalized insights but also creates vulnerabilities to breaches and misuse. Unauthorized access to a patient's comprehensive medical and lifestyle history could lead to severe consequences such as identity theft, reputational harm, and illicit data exploitation. For this reason, **safeguarding privacy** is crucial for securing public trust and ethical acceptance.

Notwithstanding this, maintaining data anonymization poses specific challenges under frameworks like the GDPR and the European Health Data Space. **Balancing** the need for anonymized data while **preserving the personalized utility of Digital Twins** remains a complex task.

VHT also raises nuanced issues of data ownership and informed consent. Patients often lack full understanding of the scope of data collection or its potential repurposing for secondary uses like research or policymaking. Ethical principles necessitate transparency toward patients, requiring clear

explanations of data collection, processing, and usage to enable informed decisions. Furthermore, long-term data usage introduces complications, as patients' preferences may evolve. Adaptive consent models, such as dynamic or broad consent, are recommended to address this challenge, empowering patients to adjust their data-sharing preferences over time.

26.1.2 Accuracy and misrepresentation

The reliability of VHT hinges on their **ability to accurately represent a patient's health status**. Errors or outdated data can lead to misdiagnoses, inappropriate treatments, or adverse outcomes. The data informing these models are inherently complex, encompassing physiological metrics, psychological assessments, and emotional states, collected through diverse sources such as sensors and algorithms. Regular DT updates are essential to maintain the models' validity and relevance.

Patients' rights under data protection laws, such as the GDPR, allow them to update, withdraw, or delete their data. While these rights are critical for individual autonomy, they can impact model accuracy. The ethical deployment of VHT requires **rigorous validation processes** to minimize errors and ensure accurate health representations.

Another challenge lies in the reliance on data from wearables and wellness Apps, which often lack the clinical-grade precision required for medical decision-making. These devices, although accessible, risk introducing inaccuracies that could distort a patient's digital representation. Therefore, developers must carefully evaluate the quality of data from such sources to meet ethical and legal standards.

26.1.3 Ownership and control

Ownership of data within VHT is a complex and multifaceted issue, involving questions of control, rights, and responsibilities. Patients, as the primary data sources, often feel entitled to ownership, aligning with principles of personal autonomy and privacy. However, healthcare providers, researchers, and technology developers may claim ownership based on their investments in building and maintaining the VHT infrastructure.

Conflicts can arise when patients seek to modify or delete data that healthcare providers believe are critical for clinical accuracy. Additionally, questions emerge around derivative insights generated by VHT, such as new research patterns or predictive markers. Determining ownership of these insights – whether they belong to patients, researchers, or healthcare institutions – presents ethical challenges.

Third-party involvement further complicates matters. For example, insurers or pharmaceutical companies may seek access to VHT data for tailoring services or evaluating risks. Patients often fear exploitation or misuse, such as discriminatory practices in insurance pricing. Legal protections are necessary to ensure that VHT data are not misused, particularly in non-clinical contexts, safeguarding patient rights and preventing ethical breaches.

26.2 Autonomy, fairness, and the risks of over-reliance on VHT technology

26.2.1 Patient autonomy

Patient autonomy, rooted in the right to informed, voluntary healthcare decisions, assumes new dimensions in the context of VHT. These **technologies extend autonomy beyond physical health** decisions to include control over digital representations of health profiles. Informed consent becomes more complex as patients must understand not only direct interventions but also the continuous use of their data in predictive analytics.

Patients may face subtle **pressures to follow VHT-generated recommendations**, potentially compromising their independent decision-making. For instance, a VHT's prediction of high-risk conditions might compel patients to accept certain treatments, undermining their freedom to choose alternatives. Similarly, healthcare providers might prioritize Digital Twins-driven insights over patient-reported symptoms, leading to ethical tensions. Transparent communication is essential to preserve autonomy, ensuring that patients are well-informed about the use and implications of their personal and sensitive data.

Additionally, respecting autonomy requires **balancing patient preferences with broader ethical principles like beneficence and non-maleficence**. Patients should have the right to decline predictive interventions, even when data-driven recommendations suggest otherwise. The ethical use of VHT

demands a careful equilibrium between respecting individual choices and leveraging the technology's predictive capabilities.

26.2.2 *Fairness and equality*

The equitable distribution of VHT benefits is a pressing ethical concern. The advanced infrastructure and resources required for VHT **deployment often limit access to wealthier institutions and individuals**, risking the creation of a healthcare divide. Underserved communities may miss out on the advantages of personalized, data-driven treatments due to financial, cultural, or geographical barriers. Algorithmic biases in VHT systems also pose risks of discrimination. Models trained on homogenous datasets may **fail to represent the health realities of diverse populations**, leading to biased predictions and inequitable outcomes. Rigorous testing for algorithmic fairness and the use of diverse datasets are essential to mitigate these risks.

Transparency in VHT decision-making is equally important. Patients and providers must understand the role of these models in care, including their limitations. Open communication fosters trust and ensures that all stakeholders can make informed decisions, promoting both fairness and autonomy in healthcare.

26.2.3 *Managing technological over-dependence*

The powerful capabilities of VHT can **foster undue reliance on technology**, overshadowing human judgment and relational aspects of care. While VHT enhances efficiency and predictive accuracy, they risk reducing healthcare to data-driven processes, potentially neglecting empathy and individualized treatment.

Clinicians may feel pressured to conform to VHT recommendations, fearing deviation might be perceived as unscientific. However, **over-reliance can lead to errors**, particularly in complex or atypical cases where algorithmic outputs lack contextual understanding. Ethical frameworks like the GDPR and particularly the AI Act emphasize the **need for human oversight**, ensuring that clinicians can critically evaluate and intervene in VHT-influenced decisions.

Balancing data-driven insights with traditional care approaches is crucial for preserving the human elements of healthcare. Ethical guidelines should advocate for the **adoption of integrated use**, where technology complements but does not replace clinical expertise.

26.3 *Multidisciplinary approach to address ethical challenges*

The definition of problems or challenges inherent to its field of application can help to anticipate and reduce its potential impact. In medicine, the Digital Twin offers the potential for promising advances, enabling a better understanding of how the human body functions, promoting the development of research to improve patient health, and improving health diagnostics or medical interventions. Emerging questions about the responsible development and deployment of VHT technology have been discussed in the previous sections. Besides those questions, there are additional challenges related to the VHT that require further thought and investment, such as identity and human enhancement, in the spirit expressed by the quote “*We shape our tools and, thereafter, our tools shape us*”⁴²¹.

Ethical reflections on the VHT require collaboration between disciplines, philosophers, social and human scientists, lawyers, mathematicians, computer scientists, and representatives of society, not only patients but also healthy people.

26.3.1 *The principle of identity*

Should the Digital Twin be an **extension of a person or as an asset in an immaterial representation of the individual**? Indeed, the Digital Twin is the creation of digital replicas of individuals, based on their data. We have here a hybrid identity as it can be a person or an object or a way of interpretation of data. This challenges the legal system and questions principles such as Identity, free will and property. As a consequence, legal scholars urged to remain vigilant of legal risks from Digital Twin projects. The notion of digital human rights is emerging and Marina Teller, lawyer, questions the concept of person, identity, entitlement to rights and obligations, legal capacity, liability, data

⁴²¹ Quote from J.M. Culklin, in the spirit of M. McLuhan. Culklin, J.M. A schoolman's guide to Marshall McLuhan. Saturday Review, pp. 51-53, 71-72, 1967.

processing (Teller, 2021⁴²²). She questions personhood and ownership as in the digital age, individuals become collections of data points, raising concerns about how law should address this “derealisation” of personhood. Marina Teller highlights the philosophical questions behind “*The line between human beings and things blurs, and the twin may occupy a legal grey area as a center of interest, neither fully a person, nor merely an object. It blurs the line between being and having.*”

26.3.2 Ethics of human enhancement

In addition to the blurring lines between being and having, the **distinction between therapy, preventative care, and enhancement can also become blurred**. When a Digital Twin approach would be applied to health care, a shift in related concepts can be expected. The already problematic distinction between therapy and enhancement becomes more critical when taking the individual’s normal patterns as reference in a DT approach. Therapy entails the maintenance or restoration of this individualized normal state, so that this distinction contains an important normative element to consider. We need to consider that concerns that were raised in the context of genomics will be even more relevant in the case of VHT applications, since their combination of multiple layers of biological (possibly including genomics) and behavioural data will be much more telling about a person than genomics data alone. “*A DT for a human may be not just a powerful tool to improve one’s physical condition. It may also be a second self who can – metaphorically speaking – rise up against its biological counterpart. [...] In this way, it may be the case that the only way to achieve equality of capabilities would be by creating data which may in turn be used to penalize some groups or to create new forms of discrimination.*”⁴²³

26.3.3 Ethics-by-design

The current revolution is based above all on a technicist ideology, and philosopher Vanessa Nurock stresses the need for philosophers to be involved in ethical and political reflections on this powerful technology, and not to leave it to engineers and computer scientists alone to propose new guiding principles in the social, ethical and political fields suggesting thus a **multidisciplinary collaboration** involving a range of skills (Nurock, 2024⁴²⁴). Nurock also stresses the importance of **situating ethical reflection** not only downstream of social experimentation with AI, but **throughout the development process**, focusing not only on principles but also on consequences. While philosophy can address downstream issues, the concept of Ethics by Design (Nurock et al., 2021⁴²⁵, Brey and Dainow, 2023⁴²⁶) seems essential to help in the development of trusted AI. **Ethics-by-Design** is an approach that enables ethical considerations to be systematically and comprehensively included in the design and development process of new technological systems and technologies. The ethics guidelines for trustworthy AI, developed by the High-Level Expert Group on AI, have been updated to incorporate concrete technological advice for overcoming challenges identified by an AI system.

⁴²² Teller M., 2021, Legal aspects related to digital twin Phil. Trans. R. Soc. A. 37920210023 <http://doi.org/10.1098/rsta.2021.0023>

⁴²³ <https://doi.org/10.3389/fgene.2018.00031>

⁴²⁴ Nurock V., 2024, Journal of Artificial Intelligence for Sustainable Development, Vol. 1, No. 1 (2024), 3–8, <https://doi.org/10.69828/4d4kf7>

⁴²⁵ Nurock V, Chatila R, Parizeau M-H. What does ‘ethical by DESIGN’ mean? In: Braunschweig B, Ghallab M, editors. Reflections on Artificial Intelligence for Humanity. Cham, Switzerland: Springer International Publishing; 2021. pp. 171–190. doi: 10.1007/978-3-030-69128-8_11

⁴²⁶ Brey, P., Dainow, B. Ethics by design for artificial intelligence. AI Ethics (2023). <https://doi.org/10.1007/s43681-023-00330-4>

27 The social impact of the VHT

This chapter discusses activities and the resulting discussion points organised to assess the social impact of the VHT. Whereas some elements overlap with the previous chapters on legal and ethical aspects of the VHT, they are still in order to properly reflect the concerns and questions raised by the diverse stakeholders participating in the activities.

27.1 Responsible Research and Innovation

27.1.1 From ELSI to RRI

Besides ethical and legal considerations, it is paramount to study and address the **social dimensions** of new technologies, including the VHT, and the inevitable impact on the society they will be embedded in^{427,428,429}. The VHT promises to change modern healthcare as we know it, by reducing costs, increasing effectiveness, and enabling personalized medicine⁴³⁰. The successful uptake of technologies from the scientific sphere into society necessitates **collectively navigating societal challenges** as they present themselves. Although societal deliberation on scientific advancements has a long history, the 1990s marked a shift in European institutions' efforts to govern public trust in science. With this emerged formal frameworks to address and manage societal and ethical implications of science, such as the introduction of the **ELSI program** in the US as part of the Human Genome Project⁴³¹. A few decades later, the European Commission introduced a renewed method to conduct research and innovation in an ethically responsible and socially beneficial way, called the **Responsible Research and Innovation (RRI)** framework. This move marked a semantic shift from ELSI, broadening its emphasis on the socio-economic impact of research and innovation (R&I).

Introduced as part of the Horizon 2020 funding scheme, RRI aims to better **bridge the gap between those who make science and those who reflect on it**, by including a variety of stakeholders, including non-technical ones, in all steps of R&I.⁴³² An ideal RRI framework, as articulated by Stilgoe and colleagues (2013), should anticipate future scenarios, whilst reflecting on the impact of R&I through the inclusion of diverse stakeholders, to better respond to the needs, concerns, and values expressed by these stakeholders. Proper implementation of RRI should thus ideally result in more engaged public and reflexive researchers⁴³³.

Oftentimes, however, the framework gets oversold in proposals as the magical fix to society's grand challenges. Nonetheless, Science, Technology and Society (STS) scholars⁴³⁴ have highlighted interesting counterarguments to this claim, such as **RRI often being reduced to a tick-box exercise**, resulting in a low implementation of meaningful RRI, the *nefast* impact of "projectification", the neglect of unequal power dynamics and politics, the greenwashing of technologies, and the prioritization of outreach and consensus-building over genuine collaboration^{435,436,437,438}. It is crucial to acknowledge these potential shortcomings to realize implementing RRI in practice demands significant attention. Nonetheless, the framework advocates for long-standing interdisciplinary collaborations, which makes it a valuable guiding framework for the realization of VHTs, as endorsed by different scholars working in RRI and voiced by people working in the *in silico* field^{439,440}.

⁴²⁷ Aarden, E. (2016). <https://doi.org/10.1016/j.atg.2016.09.001>

⁴²⁸ Iqbal, J. D., Krauthammer, M., & Biller-Andorno, N. (2022). <https://doi.org/10.1017/jme.2022.97>

⁴²⁹ Zwart, H., Landeweerd, L., & van Rooij, A. (2014). <https://doi.org/10.1186/s40504-014-0009-3>

⁴³⁰ Popa, E. O., van Hilten, M., Oosterkamp, E., & Bogaardt, M. J. (2021). <https://doi.org/10.1186/s40504-021-00113-x>

⁴³¹ Braun, M., & Müller, R. (2024). <https://doi.org/10.1007/s00146-024-01986-0>

⁴³² Stilgoe, J., Owen, R., & Macnaghten, P. (2013). <https://doi.org/10.1016/j.respol.2013.05.008>.

⁴³³ Felt, U., Öchsner, S., Rae, R., & Osipova, E. (2023). <https://doi.org/10.1080/23299460.2023.2235931>

⁴³⁴ This refers to the broad group of scholars critically researching the interaction between science, technology and society.

⁴³⁵ Tabarés, R., Loeber, A., Nieminen, M., Bernstein, M. J., *et al.* (2022). <https://doi.org/10.1080/23299460.2022.2101211>

⁴³⁶ van Oudheusden, M. (2014). <https://doi.org/10.1080/23299460.2014.882097>

⁴³⁷ Smolka, M., Doeze, T., & van Schomberg, L. (2024). <https://doi.org/10.1080/23299460.2024.2373922>

⁴³⁸ Felt, U., Öchsner, S., Rae, R., & Osipova, E. (2023). <https://doi.org/10.1080/23299460.2023.2235931>

⁴³⁹ Bruynseels, K., de Sio, F. S., & van den Hoven, J. (2018). <https://doi.org/10.3389/fgene.2018.00031>

⁴⁴⁰ Popa, E. O., van Hilten, M., Oosterkamp, E., & Bogaardt, M. J. (2021). <https://doi.org/10.1186/s40504-021-00113-x>

27.1.2 Stakeholder engagement efforts

In the last years, important collaborative efforts have been undertaken in the *in silico* medicine and DT community, aiming to shed light on the social impact of DTs through the inclusion of non-technical stakeholders, while managing societal expectations. Different EU-funded projects⁴⁴¹ organized multi-stakeholder focus groups in different European countries (Belgium, Italy, Hungary, France, UK), inviting a broad range of stakeholders (including, but not limited to clinicians, patients, legal, social, and ethical experts) to hear their concerns, values, and expectations related to *in silico* medicine⁴⁴². These projects involved social science partners with expertise in stakeholder engagement. In addition to focus groups with stakeholders, reflexive workshop were organised for the consortium members to further discuss the different social implications and collaboratively reflect on these topics⁴⁴³. Key topics emerging from those discussions, unsurprisingly, include ‘trust’ and ‘responsibility’, well-known buzzwords⁴⁴⁴, but nevertheless valuable within the community. Based on those results, several interactive breakout sessions were organised during the EDITH-CSA public meeting in Paris^{445,446}.

The table below summarizes the main elements derived from the aforementioned activities, complemented with social science literature on the social impact of Digital Twins, CM&S, AI, and machine learning^{447,448,449,450,451}.

Table 5: Overview of identified social implications

Benefits	Implications
Reduced need for animal and clinical testing	Data concerns (quality, availability, (re)use, transparency, access, protection, commodification, interoperability)
Cost-reduction	Environmental costs
Synthetic data generation to mitigate bias	Bias in input and output
Increase patient empowerment and autonomy	Impact on doctor-patient relationship and possible strain on healthcare system
Enable personalized healthcare	Unequal access, discrimination and stigmatization
Faster drug development	Definitional unclarities (VHT/ <i>in silico</i> medicine/ AI)
Improved treatment outcomes	Blurred boundaries healthy vs unhealthy (role of epigenetics)
Scalability	Impact on traditional notion of expertise
Sustainability	Responsibilities
Risk reduction	Trust and social acceptance

The following sections address the social implications, including concerns around data (quality, transparency, access, protection, commodification), cost and environmental impacts, effects on doctor-patient relationships and potential strain on healthcare systems, unequal access and possible discrimination, definitional ambiguities surrounding emerging technologies, blurred distinctions between health and illness, evolving notions of expertise, unclear responsibilities, and the topics of trust and social acceptance.

⁴⁴¹ For more information, see <https://www.simcor-h2020.eu/>; <http://insilico.world/>; <https://www.simcardiotest.eu>

⁴⁴² Elhadj, E., Van Horenbeeck, Z., Lievevrouw, E., Geris, L., and Van Hoyweghen, I. (2023) <https://lirias.kuleuven.be/retrieve/740366>.

⁴⁴³ Elhadj, E., Van Horenbeeck, Z., Lievevrouw, E., Geris, L., and Van Hoyweghen, I. (2023) <https://lirias.kuleuven.be/retrieve/740365>.

⁴⁴⁴ Bensaude Vincent B. (2014) <http://doi.org/10.1177/0963662513515371>

⁴⁴⁵ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH-EM-Paris_Minutes.pdf

⁴⁴⁶ After ethical approval (SMEC, KU Leuven) and informed consent, all group discussions were audio-recorded and transcribed. The pseudo-anonymized data from these engagement efforts were securely stored and iteratively coded by social scientists for qualitative data analysis using Atlas.ti software.

⁴⁴⁷ Rosemann, A., & Zhang, X. (2022). <https://doi.org/10.1016/j.jimed.2021.12.002>

⁴⁴⁸ Steerling, E., Siira, E., Nilsen, P., Svedberg, P., & Nygren, J. (2023). <https://doi.org/10.3389/frhs.2023.1211150>

⁴⁴⁹ Winter, P., & Carusi, A. (2022). <https://doi.org/10.23987/sts.102198>

⁴⁵⁰ Popa, E. O., van Hilten, M., Oosterkamp, E., & Bogaardt, M. J. (2021). <https://doi.org/10.1186/s40504-021-00113-x>

⁴⁵¹ Bruynseels, K., de Sio, F. S., & van den Hoven, J. (2018). <https://doi.org/10.3389/fgene.2018.00031>

27.2 What is a DT and who uses it?

27.2.1 The social dimension of data-related concerns

The conventional literature on the social impact of such data-driven technologies has primarily focused on data-related concerns (also discussed in the previous chapters on legal and ethical aspects), raising questions about the quality, availability, sharing, (re)use, commodification and interoperability of data^{452,453}. In theory, the VHT promises to make personalized care possible through the use of computational modelling and simulation (CM&S) techniques, allowing them to tailor outcomes and treatments to meet specific needs. The delicate balancing act between encouraging the advancement of VHTs through the (re)use of data, and safeguarding the privacy of individuals, opens up other interrelated questions, around **data access and transparency**^{454,455,456}. Without transparency on where the input data stems from and who gets access to the data, **potential biases** remain unaddressed, obscuring the implications of the models, which in turn can lead to a lack of control and autonomy over the VHT applications. If the input data to DTs are biased, the resulting outcomes could further perpetuate existing inequalities due to bad representation rather than reduce them, a scenario that repeatedly came up during group discussions, requiring consideration during the model development stage. Following this line of thought, while minorities and isolated communities could benefit greatly from these applications, they are also at a higher risk of being unable to afford to use them, a concern often cited in the literature^{457,458}.

A key plus of *in silico* medicine is its ability to **generate synthetic data**; creating digital cohorts for virtual testing to complement traditional clinical trials. This has the potential to increase the representation of underrepresented groups, such as minorities or children, who are often excluded from conventional clinical testing due to ethical constraints⁴⁵⁹. Nonetheless, even when validated and regulated for a large population, **equal access** to *in silico*-tested medical devices and the VHT **is not a given**, and should be considered during the development phase. HTA is of critical importance in this regard, as the national reimbursement strategies formulated based on its assessment determine public access to these technologies. For HTA decisions to be effective and inclusive, patient-reported outcomes are to be considered, highlighting the importance of Public and Patient involvement (PPI) early on⁴⁶⁰. For an inclusive development of the VHT, it will be essential to bring all perspectives into the conversation, including hard-to-reach and often underrepresented groups, a challenge that will require close collaboration with social sciences and humanities^{461,462}.

27.2.2 Roles of clinicians and patients

A key theme that emerged during all of the focus groups was the role of VHTs, and the potential need to redefine traditional notions of **expertise in the age of personalized medicine**, with increasing power given to digital technologies. Questions like “Who will be the expert?” and “Who should be given the final say, if a VHT undermines a clinician’s experience?” must be addressed early in the development process. Across all deliberations – whether patients, clinicians, or modelers – there seemed to be unanimous agreement that the **clinician should retain the final decision**, reflecting a shared concern over the prospect of a full takeover by digital tools. The complementing role of the VHT, leading to a reduction of animal and clinical testing, rather than full replacement was widely welcomed by most participants⁴⁶³. This reflects frequently cited concerns around the impact of disruptive technologies like the VHT on the labour market⁴⁶⁴. Less consensus, however, was obvious regarding the ideal **role patients should play** in the decision-making process, in light of digitalization of healthcare. Should

⁴⁵² Cordeiro, J. v. (2021). <https://doi.org/10.3389/fmed.2021.647897>

⁴⁵³ Rosemann, A., & Zhang, X. (2022). <https://doi.org/10.1016/j.jimed.2021.12.002>

⁴⁵⁴ Huang, P. H., Kim, K. H., & Schermer, M. (2022). <https://doi.org/10.2196/33081>

⁴⁵⁵ Iqbal, J. D., Krauthammer, M., & Biller-Andorno, N. (2022). <https://doi.org/10.1017/jme.2022.97>

⁴⁵⁶ Leo, C. G., Tumolo, M. R., Sabina, S., Colella, R., Recchia, V., *et al.* (2022). <https://doi.org/10.3390/ijerph19031510>

⁴⁵⁷ Cordeiro, J. v. (2021). <https://doi.org/10.3389/fmed.2021.647897>

⁴⁵⁸ Popa, E. O., van Hilten, M., Oosterkamp, E., & Bogaardt, M. J. (2021). <https://doi.org/10.1186/s40504-021-00113-x>

⁴⁵⁹ Badano, A., Lago, M., Sizikova, E., Delfino, J., Guan, S., Anastasio, M., & Sahiner, B. (2023). <http://arxiv.org/abs/2301.08719>

⁴⁶⁰ <https://www.eu-patient.eu/globalassets/projects/hta/hta-epf-final-report2013.pdf>

⁴⁶¹ <https://falling-walls.com/sites/default/files/medien/3/dokumente/fw-engage-perspective-review-engaging-the-excluded.pdf>

⁴⁶² Felt, U. (2014). <https://doi.org/10.1080/09505431.2014.926146>

⁴⁶³ Elhadj, E., Van Horenbeeck, Z., Lievevrouw, E., Geris, L., and Van Hoyweghen, I. (2023) <https://lirias.kuleuven.be/retrieve/740366>.

⁴⁶⁴ Rosemann, A., & Zhang, X. (2022). <https://doi.org/10.1016/j.jimed.2021.12.002>

patients remain mere passive data sources, or should a more active and participatory role be envisioned for them? These societal considerations should be considered and patients should partake in the development of the VHT from the get-go. The VHT promises to empower individual patients by giving them access to their VHTs, enabling them to actively engage in their healthcare journey. However, this **empowerment** has a flip side as it could, as addressed in the discussions, burden the already strained healthcare professionals through the anticipatory generation of health risks resulting in increased demand for preventive procedures⁴⁶⁵. Also, empowered patients might furthermore challenge medical authority, circling back to the previously discussed concern of shifting expertise⁴⁶⁶.

27.2.3 Further clarifications

A substantial portion of the discussions with patients and clinicians included **definitional clarifications** around what *in silico* medicine exactly is and does, how it differs from – more known – AI technologies, and the specific role of models in the VHT. This led to often intense discussions about the input factors for these models – who decides which factors get included/excluded? Given modelers' often-admitted limited awareness of ethical and social considerations, how can we then ensure the 'right' factors get incorporated into their work⁴⁶⁷? These considerations should be considered early on. On top of that, how will the VHT impact boundaries between healthy versus sick, and therapy versus enhancement⁴⁶⁸? How comprehensive can, and do we want, the VHT to be? Can the VHT account for the complexities of the human body, and should it? These questions are raised in the discussions and should be addressed properly by the VHT ecosystem.

As stated before, personalized medicine, enabled through the use of VHTs, would not be possible without huge amounts of data to feed the computer models. These vast quantities of data and the computational power required for such models to run also come with significant environmental costs, as they demand considerable energy resources⁴⁶⁹. Therefore, the **environmental impact** of the VHT should not be ignored, but deserves further scholarly attention. Finally, literature has come to shed light on the interesting **insider-outsider paradox** in the context of AI technologies in healthcare⁴⁷⁰; those on the outside of R&I, such as laypeople, may not fully understand the complexities of these models due to a lack of digital literacy and technical knowledge. Conversely, the ones directly involved in R&I, such as technical experts, often lack awareness of, or simply the capability to consider, the social and ethical implications of their work. Interactions with the modelling community shows a growing desire for training on the ethical, legal, and social aspects of their work. For these and many other reasons, close collaboration between social scientists and humanities, technical experts, and a representative portion of the public will be essential. The socially and ethically responsible realization of the VHT should be a shared community effort.

27.3 The importance of earning trust for the adoption of the VHT

27.3.1 Key drivers of trust

Trust, a true buzzword with the rise of digital health technologies, is currently under scrutiny within the *in silico* community. Trust in the VHT, like any other novel technology, is considered essential for successful adoption in society, and researchers are often eager to identify the factors complying with one's trust. This line of reasoning risks oversimplifying the concept's meaning, and implies that science should automatically be trusted, not contested⁴⁷¹. Instead, we should **ask what placing 'trust' in disruptive technologies means for different stakeholders**. Key drivers of trust in digital technologies, as backed up by literature, include **validation, credibility, explainability, transparency** in both development and communication, and **level of awareness**^{472,473}. In line with research on trust in digital

⁴⁶⁵ Sauerbrei, A., Kerasidou, A., Lucivero, F., & Hallowell, N. (2023). <https://doi.org/10.1186/s12911-023-02162-y>

⁴⁶⁶ Lievevrouw, E., & Van Hoyweghen, I. (2019). The social implications of digital health technology. ISBN: 9789082807011

⁴⁶⁷ Bak, M. A. R. (2022). <https://doi.org/10.1177/0037549720932656>

⁴⁶⁸ Lievevrouw, E., & Van Hoyweghen, I. (2019). The social implications of digital health technology. ISBN: 9789082807011

⁴⁶⁹ Popa, E. O., van Hilten, M., Oosterkamp, E., & Bogaardt, M. J. (2021). <https://doi.org/10.1186/s40504-021-00113-x>

⁴⁷⁰ Bak, M. A. R. (2022). <https://doi.org/10.1177/0037549720932656>

⁴⁷¹ O'Doherty, K. C. (2023). <https://doi.org/10.1080/23299460.2022.2091311>

⁴⁷² Coveney, P. v., & Highfield, R. R. (2021). <https://doi.org/10.1098/rsta.2020.0067>

⁴⁷³ Whyte, K. P., & Crease, R. P. (2010). <https://doi.org/10.1007/s11229-010-9786-3>

health technologies, the importance of keeping the ‘human’ in the loop was emphasized. Trust should be placed in human actors, not in technologies as such, highlighting the **importance of human oversight**⁴⁷⁴. Following this, trust seems to be placed in clinicians in particular and in the healthcare system in general. Healthcare professionals are key (dis)enablers in one’s willingness to engage with new technologies – provided a strong patient-clinician relationship⁴⁷⁵. While digital health technologies, like the VHT, can generate predictive health outcomes, the interpretation relies mainly on human judgement.

27.3.2 Literacy and awareness building

A strong emphasis was placed on the need for **literacy and capacity-building** and communication with non-technical stakeholders. Stakeholders, including patients, modelers, and clinicians, acknowledge the importance of stakeholder inclusion throughout every stage of research and innovation, underscoring the necessity of foundational awareness-building. **Healthcare professionals** play a key role in this respect, serving as **vital bridges** between the research community and the society that stands to benefit from these innovations⁴⁷⁶. Time and effort should be allocated to training and establishing close collaborations with them throughout the development process to ensure the proper clinical uptake of the VHT. On top of that, clinicians should be provided with the necessary tools to communicate innovative procedures with their patients, such as VR technologies, for instance (discussed extensively in the next PART).

Literacy and awareness-building efforts were believed to be crucial first steps to **introduce complex topics like the VHT to non-technical stakeholders**. Various organisations within the DT and *in silico* medicine community have taken this to heart and have produced materials to **educate stakeholders** on the technology and related aspects (including ELSI). These resources include videos series⁴⁷⁷, podcasts^{478,479}, and glossaries⁴⁸⁰.

Box 49: Success story – the VPHi info kit for stakeholder engagement

Success Story: the VPHi info kit on stakeholder engagement for DT model developers

Website: <https://www.vph-institute.org/reader-in-silico-medicine-info-kit.html>

The Virtual Physiological Human Institute (VPH institute) is the international society for *in silico* medicine. In recent years, the VPH institute has conducted a diverse range of stakeholder engagement activities as part of its involvement in EU-funded research projects, including SimCardioTest, SimCor, *In silico* World, REALM, and EDITH-CSA.

The insights and experiences gained from these activities are now compiled in the *In silico* Medicine Info Kit, along with all the practical material developed to run the activities. This info kit is designed to empower researchers, project managers, and directors involved in the proposal, planning, and execution phases of *in silico* medicine-related projects. The handbook provides clear instructions, guidelines, and resources (including videos, podcasts and success stories) to navigate the complexities of stakeholder engagement effectively.

As a living document, this info kit will be regularly updated to reflect the latest advancements in the field. Contributions from the *in silico* community are highly encouraged.

⁴⁷⁴ Steerling, E., Siira, E., Nilsen, P., Svedberg, P., & Nygren, J. (2023). <https://doi.org/10.3389/frhs.2023.1211150>

⁴⁷⁵ Winter, P., & Carusi, A. (2022). <https://doi.org/10.23987/sts.102198>

⁴⁷⁶ Rhodes, R. (2020). *The Trusted Doctor: Medical Ethics and Professionalism*. New York, NY: Oxford University Press.

⁴⁷⁷ Video series available on YouTube <https://www.youtube.com/playlist?list=PLo-UtC9aT9rvmgBZAienlaGpBTHrWXxf>

⁴⁷⁸ <https://www.youtube.com/playlist?list=PLo-UtC9aT9rv3fyTuXQBqxyTlsF3IcSag> or on spotify:

<https://open.spotify.com/show/5o9a2lx8rHiRO8RDSXBQvK?si=9b89895be0fd48f9>

⁴⁷⁹ <https://www.youtube.com/@InSilicoTrialsRealImpacts>

⁴⁸⁰ <https://www.avicenna-alliance.com/glossary.html>

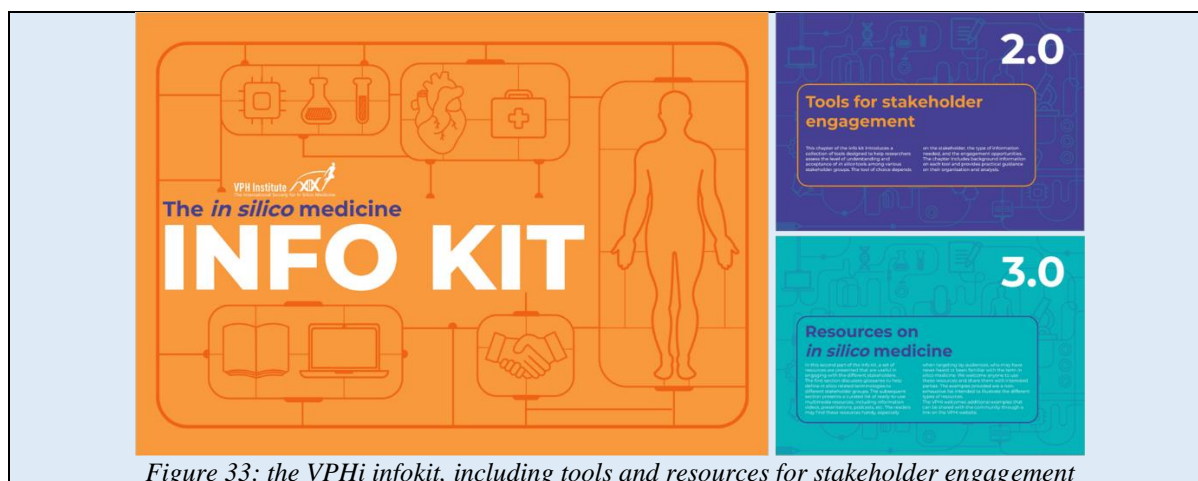


Figure 33: the VPHi infokit, including tools and resources for stakeholder engagement

27.3.3 From trust to trustworthiness

Achieving the full potential of VHT will require significant investments in engagement activities, going beyond traditional outreach efforts. Meaningful engagement is essential not only for fostering the acceptance of the technologies but also for building and maintaining meaningful relationships, understanding diverse perspectives, and promoting inclusivity and reflexivity in the development process⁴⁸¹. It is crucial to emphasize that **meaningful Patient and Public Involvement (PPI)** should not be oversold and pursued solely to build trust for a technology to reach the market; it must also aim to address and potentially dismantle power asymmetries, engaging the ‘public’ in a manner that ensures all voices are heard^{482,483}. Careful consideration should be given to what is meant by ‘the public’, whether the public we are engaging represents the public we want to reach, and the complexities of engagement, allowing space to **reflect on one’s own position and responsibilities** in this process^{484,485}. Instead of resorting to public deficit thinking and blaming a lack of perceived trust on ‘the ignorant public’, researchers working on the VHT should make space to self-reflect on what they can do to earn trustworthiness^{486,487}.

This approach shifts from assuming the public is uninformed – an externalization of one’s responsibilities – to recognizing that researchers themselves must **actively build relationships** with them and **earn trust**, which can only be done through self-reflection and engagement^{488,489}. This emphasizes that trustworthiness is something researchers must cultivate through self-reflection and meaningful engagement, rather than expecting the public to automatically trust them. As Rhodes stated (2001)⁴⁹⁰, the research community, together with healthcare personnel, should take necessary steps to ‘seek trust, and deserve it’. Interestingly, Ryan has sparked an important discourse about whether it is even possible to place trust in AI technologies⁴⁹¹. Based on three trust accounts – rational, affective, and normative – he concludes that **trusting the technology itself may be impossible**. Instead, he argues that attention should shift to evaluating the trustworthiness of the people and institutions responsible for developing these technologies, highlighting the relational aspect of trust. This **shift in focus from trust to trustworthiness** is well-studied in the domain of philosophy of science, where scholars argue trust can simply not be placed in technologies, but in the entire network behind it, from regulators to developers^{492,493}.

⁴⁸¹ Stilgoe, J., Lock, S. J., & Wilsdon, J. (2014). <https://doi.org/10.1177/0963662513518154>

⁴⁸² Wynne, B. (2006). <https://doi.org/10.1159/000092659>

⁴⁸³ Davies, S. R. (2014). <https://doi.org/10.23987/sts.55316>

⁴⁸⁴ Barnes, M., Newman, J., Knops, A., & Sullivan, H. (2003). <https://doi.org/10.1111/1467-9299.00352>

⁴⁸⁵ <https://falling-walls.com/sites/default/files/medien/3/dokumente/fw-engage-perspective-review-engaging-the-excluded.pdf>.

⁴⁸⁶ Goldenberg, M. (2021). *Vaccine Hesitancy: Public Trust, Expertise, and the War on Science*. Pittsburgh: University of Pittsburgh Press.

⁴⁸⁷ O’Doherty, K. C. (2023). <https://doi.org/10.1080/23299460.2022.2091311>

⁴⁸⁸ Wynne, B. (2006). <https://doi.org/10.1159/000092659>

⁴⁸⁹ Stilgoe, J., Lock, S. J., & Wilsdon, J. (2014). <https://doi.org/10.1177/0963662513518154>

⁴⁹⁰ Rhodes, R. (2001). <https://doi.org/10.1023/A:1014430208720>

⁴⁹¹ Ryan, M. (2020). <https://doi.org/10.1007/s11948-020-00228-y>

⁴⁹² Pink, S., Quilty, E., Grundy, J., & Hoda, R. (2024). <https://doi.org/10.1007/s00146-024-01882-7>

⁴⁹³ Winter, P., & Carusi, A. (2022). <https://doi.org/10.23987/sts.102198>

27.4 VHT and continuous stakeholder engagement

During the EDITH-CSA project, the VHT ecosystem was brought together through a range of dedicated small-scale stakeholder-specific as well as ecosystem-wide meetings and activities. As all ecosystems, the VHT ecosystem is a living, breathing thing that requires constant attention and nourishment. Hence, there is a need for meaningful public and patient involvement⁴⁹⁴, continuous stakeholder engagement, joint reflexivity and hosted public meetings. Patients, patient representatives, clinicians, modelers, policymakers, regulators, and social scientists have been given the chance to exchange ideas with a shared goal in mind: advancing health and care for all. The road to the adoption of VHT should be undertaken collectively for meaningful co-production to be possible^{495,496}, with clusters of projects joining forces (rather than repeatedly organizing similar isolated deliberative events) for more **structured, effective, and inclusive stakeholder engagement**, ensuring that all voices are heard.

⁴⁹⁴ Kiran, A. H., Oudshoorn, N., & Verbeek, P. P. (2015). <https://doi.org/10.1080/23299460.2014.992769>

⁴⁹⁵ Felt, U. (2014). <https://doi.org/10.1080/09505431.2014.926146>

⁴⁹⁶ Elhadj E., Van Horenbeeck Z., Lievevrouw E. & Van Hoyweghen I. (2024). <http://doi.org/10.1080/23299460.2024.2414484>

28 ELSI, standards and regulatory for VHT: conclusions and recommendations

28.1 Conclusions

Part 4 of the roadmap focuses on the essential ethical, legal, social, and regulatory considerations that are crucial for the responsible development and implementation of the Virtual Human Twin. This part emphasizes the need for a comprehensive framework that addresses the complex ethical, legal, and societal implications of this transformative technology.

Establishing a robust and interoperable VHT requires the adoption of common **standards** for data formats, model development, and metadata annotation. Adhering to these standards will facilitate seamless data integration across different sources, enhance model reproducibility, and promote collaboration within the VHT ecosystem. Existing standards from organizations like the International Organization for Standardization (ISO), CEN/CENELEC, and the American Society of Mechanical Engineers (ASME) provide a valuable foundation for developing VHT-specific standards, focusing on areas such as data formats, model verification and validation, metadata annotation, and terminology.

Regulatory science plays a crucial role in establishing clear guidelines and frameworks for the development, validation, and deployment of VHT technologies. A well-defined regulatory landscape will ensure the safety, efficacy, and trustworthiness of VHT applications, fostering public confidence and facilitating their integration into healthcare systems. Engaging regulatory bodies early in the development process is essential for identifying potential challenges, addressing legal and ethical considerations, and establishing a clear pathway for VHT adoption.

Health Technology Assessment (HTA) is crucial for evaluating new health technologies, including Digital Twins, to inform reimbursement decisions by payers. Current EU HTA frameworks primarily focus on evaluating digital health technologies (DHTs), encompassing solutions like digital therapeutics and clinical decision support systems. The VHT, with its personalized and predictive capabilities, aligns with these DHT categories. However, the lack of harmonized HTA methodologies and reimbursement pathways across EU member states poses a challenge for VHT adoption and reimbursement. Therefore, the VHT Community of Practice should actively engage with payers and HTA bodies to co-create an assessment framework and governance mechanisms tailored to the VHT, ensuring a streamlined approach to evaluating and reimbursing VHT solutions.

The development and implementation of the VHT raise profound **legal** questions regarding data privacy, intellectual property rights, liability, and the potential for misuse. Ensuring compliance with existing data protection regulations, such as the General Data Protection Regulation (GDPR), is paramount for protecting individual privacy and maintaining public trust. A clear legal framework addressing data ownership, access, and sharing is crucial for enabling collaborative research while respecting individual rights. Establishing mechanisms for accountability and liability in the event of unintended consequences is essential for fostering responsible VHT development and mitigating potential risks.

Ethical considerations are central to the responsible development and deployment of the VHT. Addressing potential biases in data and algorithms, ensuring transparency and explainability of VHT outcomes, and promoting equitable access to VHT benefits are crucial for preventing unintended harm and fostering public trust. Engaging ethicists, social scientists, and patient representatives in the design and governance of the VHT is essential for anticipating and mitigating potential ethical challenges.

The **social impact** of the VHT extends beyond individual patients and healthcare systems, encompassing broader societal implications. Key elements for fostering public understanding and acceptance of VHT technology, are related to building trust and mitigating potential risks of misuse or discrimination. A co-creation and stakeholder engagement process that is more than a tick-box-exercise is crucial for ensuring the responsible and equitable integration of the VHT into society. **Responsible**

Research and Innovation (RRI) principles provide a valuable framework for guiding the development and implementation of the VHT .

28.2 Recommendations

Integrating standardization, regulatory, HTA, ethical, legal, and societal considerations throughout the entire VHT lifecycle, engaging diverse stakeholders in participatory design processes, and promoting open and transparent communication are essential for ensuring that the VHT aligns with societal values and can be translated to the patient. As such, following recommendations can be made.

1. **Enabling a robust regulatory landscape for VHT:** To effectively enable the efficacy, safety, trustworthiness, performance, and risk management of the VHT from its early stages of development, it is crucial to enhance the clarity of the regulatory landscape. This can be achieved through a credibility-by-design evolutionary framework that prioritizes the establishment of clear standards, approaches, tools, and techniques. Key elements include the implementation of a supportive regulatory framework, with training and verification labs, and resources to aid in the creation of technical dossiers for certification. Robust validation and verification infrastructure is critical, as is the ability to assess data and model credibility. Ensuring data accuracy for validation and providing clear information about data accuracy will be crucial for building trust and ensuring the reliability of VHT outcomes. The establishment of dedicated sandboxes are also recommended to encourage innovation and accelerate the adoption of VHT technology.
2. **Prioritising standards for a unified and interoperable VHT:** The VHT will heavily rely on standardization to guarantee interoperability. This includes well-defined metadata standards, consistent terminology, and efficient input procedures for quality assurance. Embracing ISO standards, unifying existing standards where applicable, and prioritizing standards-based interoperability will ensure the VHT's credibility and widespread adoption. Establishing best practices and consensus procedures, particularly within mature communities that lack fixed standards, will be crucial. The development of clear guidelines for reporting and development, including standard operating procedures, will promote sharing of best practices. This approach will ensure data can be seamlessly integrated and modelling results effectively validated.
3. **Navigating the legal landscape -ensuring compliance and adoption within the VHT:** To ensure the successful implementation of the VHT within the European Union, proactive monitoring and harmonisation of relevant EU and national legislation is necessary. The VHT platform should be designed to facilitate legal compliance, incorporating features and tools that help users identify and adhere to applicable regulations. This involves developing a clear accountability framework that outlines liability, provides reassurance, and establishes responsibility through guidelines and a code of conduct for VHT users. It is crucial to identify and address challenges within existing European legislation (*e.g.* AI Act, MDR, GDPR) that might hinder the certification and adoption of VHT solutions in industry and clinical settings. A unified EU-wide approach, or at the very least, careful monitoring of national laws and recognition of differentiation across Member States will be necessary to promote legal certainty and ensure seamless collaboration.
4. **IP and the VHT, a foundation for collaboration:** Establishing common ground regarding IP Rights management and the protection of trade secrets is crucial for fostering trust and collaboration among VHT stakeholders. Additionally, the implications of software license's choices should be carefully evaluated to balance the need for open access with the protection of IP, an element that is important to facilitate the uptake of the platform and its application by start-ups and larger industrial players. Licensing frameworks operating under Fair, Reasonable, and Non-Discriminatory terms, would ensure broad accessibility while compensating innovators. A comparable approach could be applied to the VHT ecosystem where specific foundational AI tools or methods are essential for interoperable and scalable systems.
5. **The role of health technology assessment in VHT - ensuring value and efficiency:** Innovation adoption instruments, like pre-commercial procurement, innovation procurement, innovation partnerships and value-based procurement, enable evidence generation, the analysis of the impact of VHT on clinical cost structures and potential reimbursement pathways. A comprehensive innovation adoption process will facilitate translation of VHT research into practice. This includes pro-actively engaging with payers to demonstrate the synergies, value and cost-effectiveness of

VHT applications across the value chain. Harmonizing approaches across VHT adopters and Member States will be key to unlocking the competitive advantage and leadership potential of the European VHT initiative and pave the way for its wider adoption, improving healthcare outcomes and promoting sustainable healthcare systems.

6. **Building trust in VHT - ethical considerations:** The future of the VHT requires sustained efforts to build trust among users, recognizing the ethical complexities of this emerging technology. While VHT offers strong potential to increase fairness and equality in healthcare, guaranteeing digital human rights and equitable access is paramount and will constantly need to be attentively pursued. Data ownership and models, including the role of data altruism, will require careful consideration to ensure responsible development and utilization of the VHT. Continued research and open discourse on the ethical challenges and principles related to VHT, such as identity, human enhancement and technology dependence, are essential. A multi-disciplinary ethics-by-design process will ensure ethical considerations are fully embedded within the design and operation of the VHT.
7. **Embedding responsible research and innovation in the VHT:** To ensure the responsible development of the Virtual Human Twin, the genuine collaboration with Social Sciences and Humanities is crucial. This will increase stakeholder inclusion and reflexivity within the ecosystem. In addition, it will provide valuable insights into the societal impact of VHT, ethical considerations, and potential challenges related to user adoption and trust. Effective communication and dissemination strategies are essential, tailored to the specific needs and interests of various stakeholder categories. Bringing together experts from diverse fields will facilitate a holistic understanding of the complex technical, social, and ethical aspects of VHT development across the ecosystem.

PART 5:

**REALISING THE VHT – USERS, UPTAKE &
SUSTAINABILITY**

29 Incentives for uptake of the VHT

29.1 Users of the VHT

Assisted by the RRI techniques discussed in the PART 4, representatives of the identified VHT actor groups defined in the PART 1, were brought together in a thriving VHT ecosystem. This ecosystem was then leveraged to identify previously discussed technical, infrastructural, standardization, regulatory and ELSI challenges, as well as formulate a coherent vision and tangible recommendation for the realization of the VHT. Throughout the different chapters of this roadmap, care was taken to address the different members of the ecosystem, in terms of roles, responsibilities and contributions, as well as describing the different (technical) services the VHT platform can provide to its users. In this chapter we will add on to the latter and focus on those aspects that can entice users to embrace the VHT vision and its infrastructure.

29.1.1 Uptake of the VHT by users: general concepts

The involvement and commitment of the ecosystem is arguably the decisive factor for the success of the VHT initiative. Community-building activities should begin outside the infrastructure and continue within it. Activities within the infrastructure must be supported by a range of dedicated tools. The adoption of the infrastructure itself, the interoperability standards necessary for building the VHT, and the consensus processes guiding its development **require significant individual effort**. This effort must be **grounded in fulfilling the needs and expectations of the community**.

To initially engage individuals and motivate their entry into the infrastructure, it will be necessary to **keep entry barriers as low as possible**. This means that when individuals can and wish to share resources such as data and models, stringent requirements like, resources must already comply with interoperability standards or include all the metadata required by VHT resources, right from the outset, should be avoided. The gap between the maturity level of incoming resources and the level required to fully integrate into the VHT must be addressed within the infrastructure, with the support of the community. This community, in turn, must have access to **tools and services that facilitate this process** (discussed in PARTS 2 and 3). Additionally, services such as moderated forums and data curation will be crucial.

Researchers and developers, whether from academia or industry, typically work with the available data formats. Converting formats and applying standard encodings are time-consuming processes that require significant domain-specific expertise. For this reason, input and output datasets from models, as well as the models themselves, are rarely subjected to these processes unless specific project constraints make them absolutely necessary.

To move increasingly toward the adoption of standard formats and encodings, synergies with other processes and initiatives will be essential. In this context, a significant push will come from the implementation of the EHDS, as the data within the data space would have already undergone some or all of the standardization process. These standardized formats will then serve as the starting point for developing models, which in turn will produce data in the same format for it to re-enter the data space. This will trigger **virtuous cycles** where more and more data would already be available in standardized formats, and models will be defined based on these same standards. This will, in turn, facilitate collaboration and synergies within the VHT infrastructure.

29.2 Incentives for academic and industrial uptake

29.2.1 Stakeholder engagement to identify incentives

During the EDITH-CSA project, several incentivization breakout sessions were organized to probe the different VHT creators of the ecosystem on their thoughts and desires, for incentives^{497,498}. These results, combined with the results of discussions with a variety of stakeholders in different settings, including the EDITH Industry Advisory Board and Advisory Group of Stakeholders, provide the basis

⁴⁹⁷ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH-DTM-Rome_Minutes.pdf

⁴⁹⁸ https://www.edith-csa.eu/wp-content/themes/edith/documents/EDITH-EM-Paris_Minutes.pdf

for this section. The incentives focus on facilitating collaboration, improving data sharing, fostering innovation, and providing tools and recognition that enhance user engagement and productivity within the VHT ecosystem.

An assessment of the possible financial and non-financial incentives for contributions to the repository can elaborate on different remunerations.

Box 50: EDITH development – the VHT industry value proposition.

EDITH development: the VHT value proposition established with the EDITH industry advisory board
Website: <https://www.edith-csa.eu/advisory-boards/>

In order to prepare an industrial VHT value proposition, a **broad range of experts** were interviewed, covering different industry branches (pharma, medical devices, software, CRO), different company sizes and maturity levels (start-up/scale-up/SME & big industry) as well as different roles fulfilled in the company by the person answering the questions (developers, executives). The exact propositions change depending on the aforementioned factors, though, there was significant consensus on the conceptual arguments.

The following (non-exhaustive) list of **high-level value propositions** of the VHT was identified:

1. Using resources made available on EDITH's infrastructure to access a large variety of data, for validating and testing the range of applicability of a specific model, and to access existing and validated models for challenging the extension applicability of the models or the capacity of refining them.
2. Developing a robust framework for easy data sharing and legitimacy/credibility of the data being used.
3. Using publicly available portions of the infrastructure as a sand box to test new developments.
4. Using the infrastructure as benchmark for technology development, by comparing existing solutions with similar or competitive models and challenging them to reproduce observed but not yet simulated situations.
5. Benefitting from progress made on technological building blocks of the Digital Twins
6. Benefitting from attaining increased ELSI clarity & certainty, also by considering the access of data the context of the European Health Data Space.
7. Facilitating the ways of linking the companies' own developments with other partners, by favouring increased collaboration and accelerated medical innovation.
8. Highlighting new commercial opportunities, beyond the main sectors on which large software provider are currently focused, facilitating the engagement in new areas like eye, ear, and kidney diseases.

29.2.2 Incentives for VHT developers

One of the key incentives offered by the VHT platform is **access to its resource-rich infrastructure**. Users can leverage this infrastructure to explore a wide variety of data and validated models, allowing them to test and validate their own models while refining and expanding their applicability. For researchers and developers, this access significantly reduces the time and effort required to obtain high-quality resources, providing a clear motivation to engage with the platform. The ability to use publicly accessible sections of the infrastructure as a sandbox for testing and benchmarking further supports innovation and accelerates the development of new solutions. Tools and services should help researchers identify and connect to the correct resources. Quality tools embedded in the VHT infrastructure should provide a clear information on the context of use, validation status and applicability of the different resources.

Data sharing is another critical area where the VHT platform provides value. By offering a robust framework for seamless data sharing that ensures credibility and compliance with regulations, such as the European Health Data Space, the platform addresses a significant barrier for many users. Services that help transform the data and simplify the complex process of data standardization and harmonization, in compliance with regulatory standards, provide strong incentives for use and sharing.

The VHT platform also creates **incentives for collaboration** by enabling organizations to link their developments with those of other partners. This feature fosters partnerships that accelerate innovation while opening up new commercial opportunities, particularly in underexplored areas such as eye, ear, and kidney diseases. For **industry** stakeholders, this presents a chance to expand their markets and engage in pioneering research, making the platform a valuable resource for driving growth and competition.

In future developments, the platform could provide the possibility of performing part of the work under **private settings**, enabling users (or groups of users) to conduct exclusive studies without external interference before transferring them to the market and/or making them public. For industrial users, this would allow to use the platform for product development without compromising their IP position. For academic users, this would allow a research group to use the **VHT as its default platform**, pushing the results to the public domain when ready.

The platform should incentivize engagement by simplifying interactions with its infrastructure through **user-friendly interfaces and automated tools**, such as annotation services, data standardization converters, and visualization tools. These features minimize the effort required to navigate the system, ensuring that users can focus on their core work. **Knowledge graphs** that index and query the VHT catalogue further support users by identifying gaps in the knowledge landscape and highlighting opportunities for new research or development.

For *academic* users, the VHT platform offers additional, tailored incentives. The implementation of digital object identifiers and a **citation-based reward system** for uploads and executions, ensures recognition of their contributions. Citations are a much-used metric for measuring impact and career advancement. Additional services, such as annotation tools and science communication support, further enhance the appeal for academic users by addressing specific needs and reducing administrative burdens.

Related to the aforementioned science communication support, the platform should provide assistance in dissemination activities, promoting proof-of-concept stories and success cases, highlighting the benefits of open science⁴⁹⁹ and collaboration. Services should be offered that help to **build compelling high-level narratives** based on the information provided by the DT and sourced through the VHT's knowledge base. These campaigns not only build trust in the platform but also encourage users to contribute their own success stories, reinforcing the value of participation in the ecosystem.

Finally, the VHT incentivises multi-stakeholder collaboration by facilitating dialogues between clinicians, modellers, patients, and other stakeholders. Engaging patient advocacy groups early in the design process ensures that the platform remains patient-centric, while fostering interdisciplinary collaboration helps bridge gaps between different professional communities. These efforts contribute to a sense of shared purpose, motivating users to work together within the VHT ecosystem.

29.3 Clinical uptake and engagement with the VHT

29.3.1 Stakeholder engagement to identify incentives

As previously discussed, as **consumers of VHT technologies**, the needs and expectations of clinical end-users should be considerations that shape the VHT and its applications all along their development cycles, through co-creation and joint development. During the EDITH-CSA project, input was collected across the pre-selected EDITH use cases to assess the extent of clinical engagement, any challenges encountered, and the stages at which clinicians and patients were involved. Direct engagement was established with clinicians and patients as end-users by organising workshops and circulating surveys⁵⁰⁰ to collect their feedback on Digital Twins and the vision for the VHT. The information gathered included understanding issues related to the adoption of Digital Twins in clinical practice and identifying critical points that need to be addressed to enhance their clinical uptake. The incentives are strategically designed to **address clinicians' priorities and practical concerns**, emphasizing clear clinical benefits, ease of integration, and meaningful collaboration during the VHT's development and implementation stages.

29.3.2 Incentives for VHT consumers

A central incentive for clinicians is the VHT's ability to **deliver clear and demonstrable clinical benefits**. To gain their trust and engagement, the VHT must address unmet clinical needs, improve

⁴⁹⁹ <https://f1000research.com/documents/10-828>

⁵⁰⁰ <https://doi.org/10.3389/fmedt.2023.1125524>

patient outcomes, reduce costs, and enhance patient care. Unmet needs can be identified through collaboration with clinical partners, analysis of clinical funding calls, and participation in relevant events and conferences. Improved patient outcomes should align with metrics already defined in clinical practice and referenced in published guidelines. Cost reductions must be thoroughly analysed and communicated, balancing the financial implications of implementing the VHT against its potential to alleviate pressures on overburdened healthcare systems. Patient care enhancements, such as better treatment experiences, increased understanding of treatment options, and improved quality of life, further reinforce the VHT's value to clinicians and their patients. These benefits are essential motivators for clinical end-users, as they directly align with their commitment to improving healthcare delivery.

The VHT platform also **addresses practical challenges** that could deter clinicians from engaging. Ethical and legal issues surrounding data use, privacy, and sharing are critical considerations, and the VHT initiative must ensure robust mechanisms to address these concerns. The cost implications of integrating, maintaining, and/or storing VHT models within hospitals must be clearly communicated and justified in terms of their value to clinical practice. Additionally, the VHT must provide rigorous validation, verification, and regulation of its models, ensuring they meet clinical standards for reliability and safety.

Clinicians highlighted the importance of **social and economic impacts, transparency in model processes, and alignment with clinical workflows**. The latter can be realised through a variety of services directly implemented in the VHT infrastructure (*e.g.*, clinical reporting, seamless inclusion of hospital's digital workflow systems or edge devices), or provided on top of the VHT infrastructure (as third party services). By addressing these factors, the VHT can build trust and demonstrate its relevance in real-world healthcare settings. Clinicians recommended **focusing initially on small-scale, validated models that address immediate clinical needs**, providing **low-risk, high-impact demonstrations** of the VHT's potential. Representation of VHT activities and outcomes at clinical conferences and events, along with advertisements and endorsements by clinical champions would further increase trust and interest in the VHT initiative. In addition, clinician involvement at leadership levels within VHT ecosystem activities would encourage wider participation by fostering trust and ownership of the initiative.

The importance of **education and outreach** was repeatedly emphasized as a mechanism to incentivize clinician involvement. Events, workshops, and educational materials can help train clinicians, patients and the public at large, fostering a better understanding of the VHT's potential and addressing concerns about the adoption of new technologies. Ensuring that patients are engaged and involved during the development process would not only improve model applicability, but also incentivize clinicians and hospitals to participate. Tools that **facilitate patient engagement** in their care processes, are particularly appealing as they align with the broader goals of patient-centred care. Finally, **transparency in the ownership** and access rights of Digital Twins is essential to gaining patient trust, which, in turn, incentivizes clinicians to engage with the VHT.

By prioritizing these incentives and addressing the practical concerns of clinical end-users, the VHT initiative can effectively foster collaboration and accelerate its integration into clinical practice, ultimately improving patient outcomes and care delivery.

Box 51: Success story – the virtuous cycle of a partnership between Digital Twin creators and clinicians

Success story: Digital Twins for planning partial liver resection during tumour treatment

Status: from basic research (TRL 3) to validation in patients

Website: <https://team.inria.fr/simbiotx/> and <https://mimesis.inria.fr/>

The scientific partnership of computational modelling experts (I. Vignon-Clementel and team, Inria, France) and liver surgeons interested in new avenues to improve patient care (E. Vibert and team, AP-HP, Inserm & U.Paris Saclay, France), evolved over the last ten years. Their first mathematical model, parameterized from pig experimental data enabled to understand the key haemodynamic changes happening during partial liver

resection⁵⁰¹. This paved the way for creating virtual human twins in order to assess the hypertension risk of such surgery, which was **validated in a first patient cohort**⁵⁰². However, there are still many steps to achieve for it to be truly transferred to the clinics (TRL=4). Such virtual twins provide a framework to analyse the intricate hemodynamic changes involved in the novel surgery: ‘RAPID’ technique (Resection and Partial Liver Transplantation with Delayed Total Hepatectomy), a two-stage surgical process, offering insights into optimizing perfusion and graft size⁵⁰³ (TRL=3). The virtual human twin can be created in real time, thanks to machine learning⁵⁰⁴.

Another challenge in partial hepatectomy for tumour treatment is to localize tumours during the surgery, when the liver has deformed compared to preoperative imaging. This challenge has been faced by another computational modelling group (S. Cotin and team, Inria, France) and the surgeon team above (E. Vibert and team, AP-HP, Inserm & U.Paris Saclay, France). They compute the real-time deformation of the liver tissue and tumours based on the position of the liver surface, as provided by intra-operative cameras⁵⁰⁵. This augmented reality tool was **validated in vivo in four patients** and more thoroughly in an ex-vivo liver⁵⁰⁶ (TRL=4). This work is being transferred to the clinics through the **start-up** Twinical.

These patient-specific simulations, which can run real time highlight the potential of virtual twins to enhance clinical decision-making and outcomes in complex liver procedures, bridging the gap between computational modelling and translational medicine.

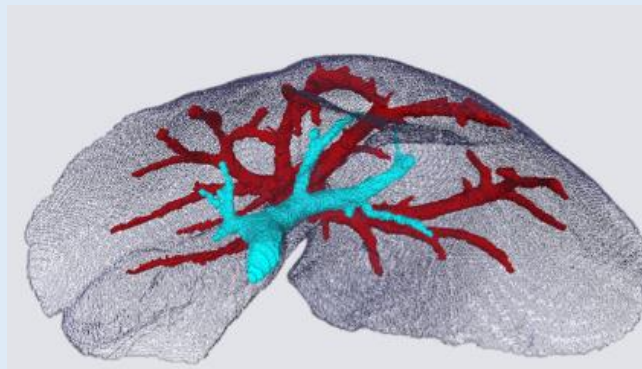


Figure 34: Digital Twin of liver tissue and vessels

Box 52: Success story – ICU Digital Twins, from Research to Clinical Application

Success story: ICU Digital Twins, from Research to Clinical Application

Status: in clinical use

Website: <https://www.clinomic.ai/>

The integration of DHT machine learning Digital Twin models into clinical practice is transitioning from an ambitious vision to an achievable reality. Whereas in the ICU context there is a clear common understanding for the strong need for DT technology, many regulatory processes—addressing privacy, ethical considerations, and legal frameworks—are still evolving and set significant hurdles for the transfer of PoC solutions to clinical practice.

Uniklinik Aachen (UKA) developed a machine learning-based Digital Twin to serve as an alarm system for the Acute Respiratory Distress Syndrome (ARDS). The DT has been extensively validated on retrospective data, is containerized, and primed from the technical side for deployment. Transfer to clinics will require extensive regulatory efforts according to the MDR regulations, which requires additional funding. The initial study already provides a valuable key element to define conceptual designs of clinical IT environments enabling ML-based DT's and will serve as a core example for the EU project ICUData4Europe.

In addition, the UKA spin-off Clinomic, a 2019 startup founded by intensive care experts from UKA, is dedicated to transforming intensive and acute care through digitalization and telemedicine. Their flagship product, Mona, provides a comprehensive platform integrating real-time data visualization and DTs in order to support the practitioners at the bedside.

⁵⁰¹ Audebert et al, J Biomech 2017

⁵⁰² Golse et al, J Hepatol 2021

⁵⁰³ Golse et al, Clinical Biomech 2020;

⁵⁰⁴ Hanna et al, submitted

⁵⁰⁵ Brunet et al, MICCAI 2019

⁵⁰⁶ Golse et al, J Gastro Surg 2020;

29.4 Facilitating access to VHT technologies through AR/VR

Facilitating the access to VHT technologies is key to achieve a fully realised VHT with large scale uptake and drive innovation. This involves not only the development of sophisticated computational solutions that accurately replicate human biology and interact with real-world health data, but also **providing the possibility of accessing them in a simplified and effective manner**, transforming a 2D visualisation into a 3D space fully immersive experience. Virtual reality (VR) technologies can deliver this. VR breaks down the physical and digital barriers of a 2D visualisation and transforms it into a full **3D immersive experience** combining enriched visualisation of complex systems and data with intuitive manipulation and interaction. Empowering VHT technologies by means of VR can unlock their full potential through simulation, stimulation, and mediation. VR for VHT can open new possibilities of development prioritising cutting-edge technologies development, addressing users' needs, advancing remote and virtual access to VHT platforms, and boosting collaboration among multiple stakeholders including researchers, trainees, clinicians and patients.

29.4.1 VR & VHT to enhance training: immersive learning.

Personalised patient care can be highly complex due to the heterogeneous, often comorbid, and clinically unstable patient population. Due to the saturation of workforce resources in Europe, along with an ever-increasing patient complexity and personalization of therapies, new solutions to **expedite clinical training of personnel** are required. VR integrating VHT technologies can contribute to addressing some of these key challenges and bring positive impact on the efficacy of medical training. Highly interactive and visually rich experiences in VR allow users to easily engage with VHT models in ways that traditional methods cannot. Thus, users, especially healthcare professionals, can fully visualize 3D anatomical replica of patient-specific organs and interact with such models in a more natural way. This brings the immediate benefit of an **easier understanding** of the spatial relationship between anatomical structures, but this is only a part of the gains. As a simulation, a DT can provide a **safe way to explore new scenarios** or try out new skills that might be hard to replicate in the real world. Thus, physics-based simulations can also be accessed by VR, enabling users to explore the dynamic behaviour of biological systems under various conditions. VR environments create safe spaces for experiencing customizable clinical scenarios that can mimic real-world conditions, from surgical procedures to emergency care situations. These scenarios can be tailored to the needs of individual trainees or to import DT data, allowing for a progressive and adaptive learning experience that builds confidence and skills in a safe, controlled environment.

One of the key advantages of VR is its **multi-user capability**, which fosters **collaborative and remote training experiences**. Trainees and experts can interact within the same virtual environment, regardless of geographical constraints. This facilitates access to highly specialized resources and expertise that may otherwise be limited or unavailable. As a result, VR can crucially contribute to the democratization of VHT solutions by bridging gaps between research and clinical practice, ensuring that cutting-edge findings and innovations are seamlessly integrated into training programs.

29.4.2 VR & VHT to support procedural planning in complex medical areas.

Interventional and surgical treatments for intricate medical cases, such as those involving brain, heart and liver, demand an exceptionally detailed understanding of the patient's unique anatomy and the complex dependencies of physiological processes. The use of DTs for personalised planning enables simulation of 'what-if' scenarios to **compare different strategies** optimising them for individual treatments. Access to models and results is, however, still mainly limited to experts.

VR can lower the barriers of access by providing a more intuitive understanding of the intricate spatial relationships between anatomical structures, such as the proximity of a tumour to vital blood vessels or the electrical pathways critical to cardiac function or to analyse physiological dependencies, such as blood flow or the impact of surgical interventions on surrounding tissues, within the immersive environment. In this context, VR can **bring together data from different sources** and make them easily accessible for a more informed decision process. Regarding training, VR makes it also possible for **multidisciplinary teams** to explore together various treatment strategies, evaluate their potential outcomes, and agree on the most effective approach.

Despite the promises, the full implementation of VR and DTH technologies for planning is still in its infancy and significant effort will need to be dedicated to the development of standardized interfaces,

data types and integration within the VHT platform to deliver a paradigm shift in the planning and execution of complex medical procedures.

29.4.3 VR & VHT for improving communication with patients and their empowerment.

Effective management of patients requires engagement and empathy of multiple stakeholders. Understanding specific conditions and **effectively communicating** are a critical part of the patients' journey and contribute to their satisfaction and the long-term success of healthcare delivery.

VR and DTH technologies can transform communication with patients by making complex medical concepts more accessible and fostering meaningful interactions. Traditionally, medical explanations rely on static images or abstract descriptions, which can leave non-experts confused or overwhelmed. By providing an **immersive and visual experience**, VR and VHT enable patients to explore 3D models of their own anatomy or medical condition in real-time. This personalized visualization helps patients understand their diagnosis, treatment options, and potential outcomes more clearly. With enhanced understanding comes **patients' empowerment**. Patients who comprehend their health conditions are better equipped to participate in shared decision-making with their healthcare providers. For instance, a patient viewing the planning of surgery can explore a VR simulation of the procedure, gaining insight into its necessity, risks, and expected benefits. Similarly, during rehabilitation, patients can monitor their progress through DT-generated feedback, fostering a sense of control and motivation. Additionally, VR can help patients to rehearse real-life scenarios, such as managing mobility challenges or coping with chronic conditions, in a safe, controlled environment.

Such a proactive involvement **promotes confidence, adherence to treatment plans, and improved health outcomes**. Together, VR and DTH create an inclusive and interactive healthcare experience, prioritizing patient engagement and empowerment at every stage of care.

29.4.4 VR & VHT: an emerging opportunity for business models.

VR and VHT are two emerging technologies that have the potential to radically change markets and industries and transform the interaction among stakeholders. These technologies are only sporadically used in medicine. This field is usually slow in reacting to innovation and uptake of new technologies. However, training, planning and patients' communication are significant opportunities for **creating or accelerating business development**.

Democratizing training and introducing new tools for doctors and patients can become soon part of the daily practice fostered by advancements in hardware that are making high-performance devices more accessible. The EU agenda supporting personalized medicine with reduced medical errors and improved access to healthcare and medical education fosters an attractive ecosystem for businesses. However, risks still need to be addressed to ensure the **creation of a healthy economic environment**. These risks include, for example, high regulatory compliance costs, particularly with MDR and CE certification, significant development expenses for accurate DTs and 3D models, physiological resistance to new technologies, the necessity of clinical validation to prove efficacy, and integration challenges with existing healthcare systems.

Mitigation strategies will need to involve expedited regulatory approvals, demonstration of clinical benefits, prioritization of user-friendly designs, and seamless IT integration. In this field it is notable to record how many **startups** originate from academic institutions. This common condition therefore requires further **institutional support to transition into viable businesses**. In this context, the growth of a VHT ecosystem can contribute to the success of such ventures and the scalability of the proposed solutions in this area.

Box 53: Success story – Virtual Reality & Digital Twins in a paediatric clinical centre

Success story : Virtual Reality & Digital Twins in a paediatric clinical centre, the VhearRs case study

Status: in clinical use

Website: <http://www.VheaRts.org>

Congenital abnormalities – structural and functional defects that develop during prenatal life – affect ~6% of newborns worldwide. These conditions contribute significantly to child mortality, with >300,000 deaths occurring annually within the first month of life. Advances in paediatric surgical procedures, interventions and medical technologies have improved immediate management of structural anomalies. However, long-term

prognosis in these patients remains poor, especially in patients with complex defects who experience substantial morbidities and low quality of life. Continuous monitoring throughout the patient's life is therefore required. To assess paediatric patient conditions and plan personalized treatments, clinicians rely heavily on advanced diagnostic imaging (*i.e.* ultrasound, computed tomography, magnetic resonance imaging), routinely presented as cross-sections of the diseased organs on 2D screens. However, the complexity and uniqueness of congenital anomalies, together with their evolution as the child grows, make it challenging for clinicians to interpret these images, fully assess the 3D anatomy, and strategize optimal interventions.

DT can enable individualised simulations of such conditions for diagnosis, training, personalised treatment planning, and contribute to optimised outcomes. Despite the transformative potential of DT for healthcare delivery and paediatric patient care, there are still substantial barriers hindering the access to such technologies and their widespread adoption.

At University College London (UCL) & Great Ormond Street Hospital for Children NHS Trust (GOSH) a group with extensive experience in computer modelling of congenital diseases, particularly cardiovascular⁵⁰⁷ and craniofacial conditions⁵⁰⁸, has developed in-house a VR software, VheaRts⁵⁰⁹, to facilitate access to DT solutions. VheaRts can be used to visualize patient-specific models and simulation by means of immersive experience.

Since its conception (2019), VheaRts has been extensively used for delivering training programmes, planning of complex procedures and enhancing communication with patients and their families.

VheaRts for training. The use of VR and DT have been a key focus to enable access to resources to healthcare professionals and medical students. VheaRts include a large library of curated, patient-specific, anatomical models available within an



immersive VR environment where a series of tools facilitate the interaction with DT models⁵¹⁰. VheaRts has unique multi-users capabilities which have allowed its adoption within UCL courses and GOSH training programs. This also fostered the creation of brand-new courses which do not replicate the traditional teaching but takes into account the full capabilities of the VR and DT technologies. To date, more than 3,000 users have tried VheaRts showing outstanding appreciation for the new way to access medical contents and, at the same time, providing feedback useful for future improvement of software and hardware.

VheaRts for personalized clinical care. The use of DT models for clinical case planning at GOSH is steadily growing also thanks to easier VR access (>more than 40 cases in 2024)⁵¹¹. The request for a VR inspection for pre-operative planning is normally requested on a case-by-case basis by clinicians, for complex cases. The workflow involves the transformation of patient images into DT models which can then be imported into the VheaRts platform for visualisation and interaction. Surgeons typically explore and map out the routes and constraints of potential repairs using markers, distances and a combination of cutting tools, all useful for understanding the spatial relationships of specific cardiovascular structures such as tumours, major vessels or chambers. Preliminary results show how VR helps the understanding of depth perception and provides an easier interaction with 3D objects if compared to traditional PC experiences. As a result, operators can achieve improved confidence in their surgical plan following VR assessment when compared to using only cross-sectional imaging. These findings suggested an ongoing multicentre clinical study on the impact of VR on planning



⁵⁰⁷ Capelli C et al. Patient-specific simulations for planning treatment in congenital heart disease. *Interface Focus*. 2018

⁵⁰⁸ He KH et al. Understanding the influence of surgical parameters on craniofacial surgery outcomes: a computational study. *Royal Society Open Science*. 2024

⁵⁰⁹ <http://www.Vhearts.org>

⁵¹⁰ Capelli C et al. Vhearts: Moving from 3D Print Models to Virtual Reality for Enhancing the Anatomical Understanding of Congenital Heart Diseases. *Journal of Cardiovascular Magnetic Resonance*. 2024

⁵¹¹ Pajaziti E et al. VheaRts: Reporting a Single-Centre Experience in Developing and Implementing a Virtual Reality Application for Planning Treatment of Congenital Heart Disease. In *International Conference on Extended Reality* 2024

VheaRts for improving patient experience.

The integration of VR at GOSH focuses on how it can be used as a tool to enhance the overall clinical care. In this context, VheaRts is used as a new communication tool aimed to increase patient awareness and understanding of clinical settings, their conditions, reducing anxiety, and creating more positive associations with medical environments. A version of VheaRts dedicated to young adults has been designed for learning at all ages. Pilot test showed how both patients and their families can benefit from a VR experience by understanding more details of specific conditions. VR and DT can also be used to involve young people in research and medicine. In this sense, VheaRts is used to support communication at schools to raise awareness not only to the diseases but also to showcase the use of such technologies in medicine as a potential future career path for the students, which they may not have considered or been exposed to before.

Sustainability. It is important that the investment in VR and DT promotes sustainability and revenue generation for all the involved institutions. For VheaRts, several routes towards these targets have been explored. The pipelines for enhanced training and personalised care can be leveraged for commercial opportunity as other institutions would value the opportunity to take advantage of these modalities. The growing market involves healthcare and academic institutions in the first instance, but it can go beyond that. For a development of a commercial plan spinning out of universities, several barriers still exist, and it is crucial the institution guarantees support and guidance to plan the transition from a research project to a commercial product. In this sense, the management of intellectual property and access to seed funding need to be facilitated.



30 Business models for VHT

The private sector is an indispensable pillar in the development and implementation of the VHT, given the critical role that health industries play in innovation, R&D investment, and implementation. Private health companies are often the driving force behind adopting and scaling groundbreaking technologies, including something as ambitious as the VHT. The following are the main reasons why it is essential that the private sector remains fully integrated within the VHT ecosystem.

- **Investment capacity.** The private sector, particularly the pharmaceutical, biotechnology, and MedTech industries, is among the largest contributors to healthcare R&D globally. For example, major pharmaceutical companies invest billions annually in R&D, eventually outpacing the initial public investment which, however, is often critical for triggering the whole process. Private financial capacity is indispensable for funding and scaling VHT applications.
- **Risk-taking and agility.** Private entities are often more capable of taking risks to explore promising technologies. In a properly supported setting, they can collaborate with startups, academic institutions, and other industry players, concurring to create a dynamic environment that accelerates VHT development. Such a collaborative approach is crucial for addressing the complex challenges associated with VHT implementation.
- **Regulatory Navigation.** Private companies have extensive experience navigating complex regulatory landscapes. This expertise is vital for ensuring VHT technologies meet stringent safety and efficacy standards, facilitating faster approval and adoption.
- **Market Integration:** The private sector is the bridge between research and real-world applications. Health industries have established pipelines for integrating new technologies into clinical practice, regulatory approval pathways, and market deployment. They are uniquely positioned to commercialise VHT, creating value for patients and providers while crucially contributing to the financial sustainability of the ecosystem.
- **Demand for VHT:** Health industries are likely to be early adopters of VHT technologies. For example, pharmaceutical companies could use VHT technologies for drug development and personalised medicine. MedTech companies might incorporate VHT technologies into devices or software for diagnostics and monitoring.
- **Public-private collaboration:** Lessons from past projects like the HGP and Human Brain Project indicate that public funding alone is insufficient to sustain long-term innovation. A balanced approach, leveraging both public grants and private sector investments, can ensure sustainable growth and faster scaling.

As VHT technology evolves, fostering strong public-private partnerships will be essential to realise its full potential in revolutionising healthcare delivery and patient outcomes.

30.1 Current & future market appraisal

The global DT in healthcare market is projected to grow at a Compound Annual Growth Rate (CAGR) of 67.0% from 2023 to 2028, reaching USD 21,113.0 million in 2028 from USD 621.0 million in 2023⁵¹². This growth is attributed to factors such as increasing investments in DT technology, its expanding applications in healthcare, and the rapid pace of technological advancements. These estimates highlight several key trends that shape the market were identified, including the increasing use of DTs and the growing emphasis on patient-centric care.

30.1.1 Growth across all applications

Growth is projected to occur in all key applications for Digital Twins in healthcare (figure below), including:

- **Drug discovery & development:** Digital Twins are used to model and simulate the effects of drugs on virtual patients, accelerating the drug development process and reducing costs. This application is expected to witness the highest CAGR during the forecast period.

⁵¹² Digital twins in healthcare market: Global forecast to 2028; Market and Markets, 2024. The source is restricted and not publicly accessible. The underlying research relied on two approaches. One approach was based on an analysis of the market through written resources, including scientific literature, reports produced by relevant organisations and centres, websites, press releases as well as data repositories. Another approach was to interview key opinion leaders from the supply side (manufacturers, suppliers, distributors and vendors), the demand side (Pharma & Biopharma Companies, Hospitals, Academic Institutes, Research & Diagnostic Labs, and Medical Device Companies), as well as others (consultants, Government officials & regulatory authorities). The report examines various aspects of the market, including market segmentation, growth projections (for 2028), key drivers and restraints, regional analysis, and competitive landscape.

- **Personalised medicine:** Digital Twins enable the development of personalised diagnostics and treatment plans tailored to individual patients' genetic and clinical profiles.
- **Surgical planning & medical education:** Digital Twins allow surgeons to plan and practice complex procedures on virtual models, improving surgical outcomes and reducing risks. They also serve as valuable tools for medical education and training.
- **Medical device design & testing:** Digital Twins are used to design and test medical devices in virtual environments, reducing development time and costs.
- **Healthcare workflow optimisation & asset management:** Digital Twins can optimise healthcare workflows by simulating and analysing different scenarios, leading to improved efficiency and resource allocation. They can also be used to manage and monitor medical equipment, reducing downtime and maintenance costs.
- **Other applications:** This category includes applications such as risk adjustment, remote patient monitoring, medical imaging, and diagnostics.

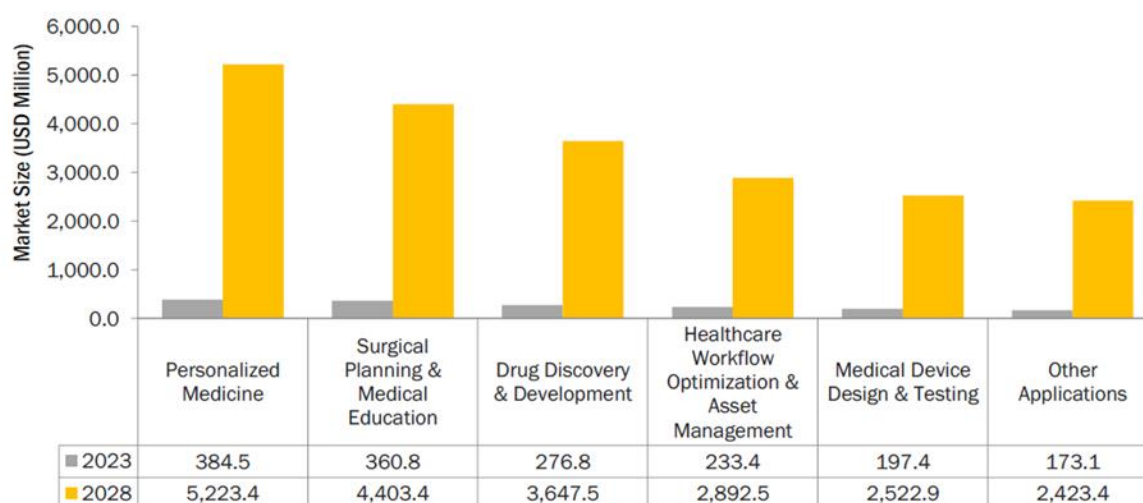


Figure 35: Digital Twins in the Healthcare market, by application, 2023 vs 2028 (USD million)⁵¹³.

30.1.2 Growth across all End Users

Growth is projected to occur for the various end users of DTs in healthcare (figure below), including:

- **Pharma & biopharma companies:** These companies are using Digital Twins to accelerate drug development, personalise treatments, and improve clinical trial efficiency.
- **Research & academia:** Digital Twins are increasingly being used in academic and research institutions for medical research, disease modelling, and drug discovery.
- **Healthcare providers:** Hospitals and clinics are adopting Digital Twins to optimise workflows, improve patient care, and manage medical equipment. They are projected to be the largest end-user segment during the forecast period.
- **Medical device companies:** Digital Twins are enabling medical device companies to develop and test new devices more efficiently and cost-effectively.
- **Other end users:** This segment includes payers, contract research organisations (CROs), and patients.

⁵¹³ MarketsandMarkets report 2023, based on the following sources: US Census Bureau, European Commission, Organisation for Economic Co-operation and Development (OECD)

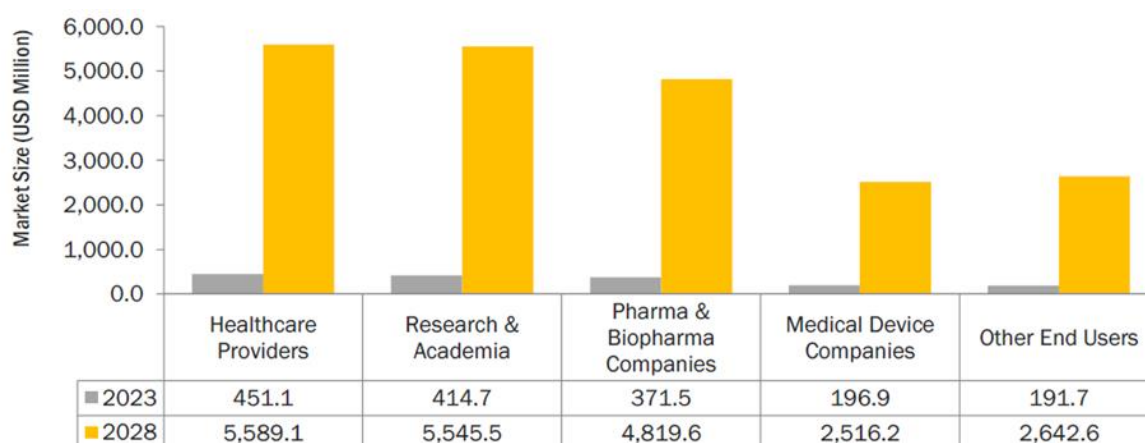


Figure 36: Digital Twins in Healthcare market, by end-user, 2023 vs 2028 (USD million)⁵¹⁴.

30.1.3 Growth across all geographical regions

The Digital Twins in the healthcare market exhibit significant regional variations in terms of market size, growth rate, and key drivers (figure below). The report focuses on four major regions: North America, Europe, Asia Pacific, and the Rest of the World.

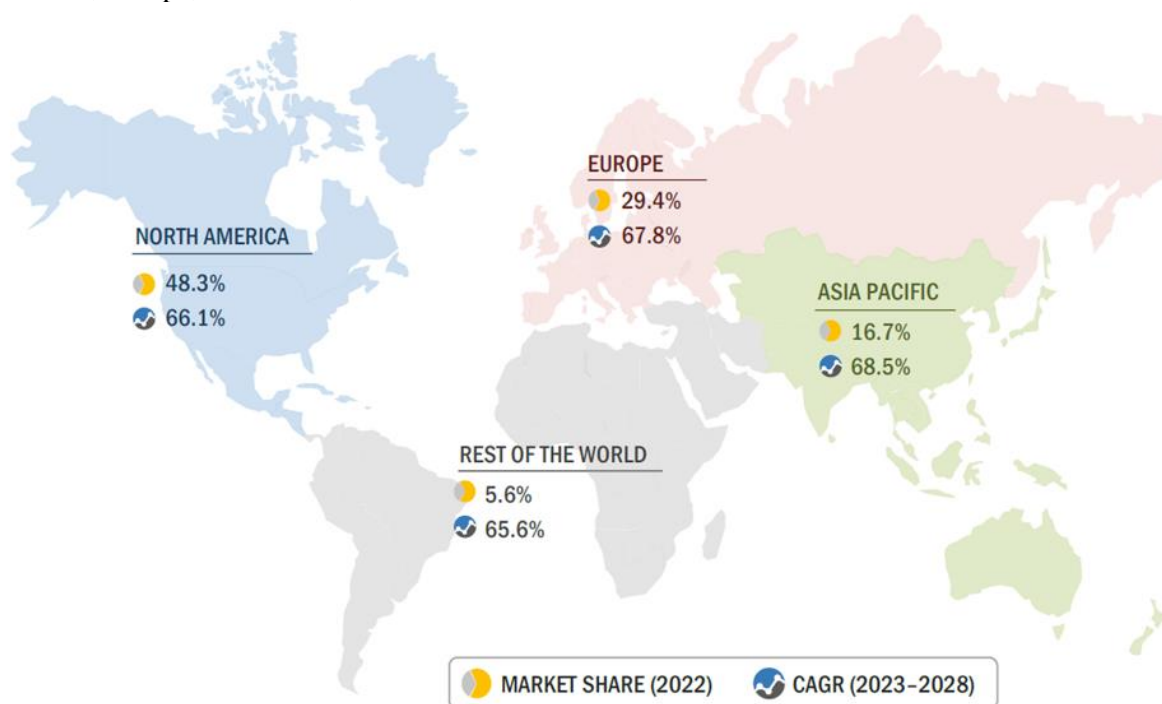


Figure 37: Geographical snapshot of Digital Twins in the Healthcare market⁵¹⁵.

Europe is the second-largest market for Digital Twins in healthcare, accounting for 29.4% of the global market in 2022. The market is projected to grow at a CAGR of 67.8% during the forecast period, reaching USD 6,393.7 million in 2028 from USD 480.4 million in 2023.

Several factors are driving the growth of Digital Twins in the healthcare market in Europe:

- **Increasing adoption of digital health technologies:** European countries are increasingly embracing digital technologies in healthcare, creating a favourable environment for the adoption of Digital Twins.

⁵¹⁴ MarketsandMarkets report 2023, based on following sources: US Census Bureau, European Commission, Organisation for Economic Co-operation and Development (OECD)

⁵¹⁵ MarketsandMarkets report 2023, based on the following sources: US Census Bureau, European Commission, Organisation for Economic Co-operation and Development (OECD), CDC, CMS, World Bank, Digital Twins Consortium, European Commission, AHA, National Center for Biotechnology Information (NCBI), NIH, WHO, Eurostat, P-MEC News, Interviews with Experts, and MarketsandMarkets Analysis.

- **Government initiatives promoting digital transformation in healthcare:** The European Commission has launched several initiatives to promote the use of digital technologies in healthcare, most recently with the aim of creating personalised healthcare solutions based on Digital Twins.
- **Growing investments in Digital Twin research & development:** European countries are investing heavily in research and development activities related to Digital Twins in healthcare, leading to technological advancements and market expansion. An example of this is the MediTwin project, a 130 mio€ investment by the French government in the framework of the France2030 initiative, which brought together 14 partners from academia, industry, and hospitals to develop and deploy Digital Twins in healthcare.
- **Presence of key market players:** Europe is home to several leading companies in the Digital Twins in the healthcare market, such as Dassault Systèmes (France), Koninklijke Philips N.V. (Netherlands) and Siemens Healthineers AG (Germany),

Within Europe, Germany is the largest market for Digital Twins in healthcare, followed by the UK and France. The market in Germany is projected to grow at a CAGR of 66.3% during the forecast period, reaching USD 1,849.3 million in 2028 from USD 145.3 million in 2023. The UK market is expected to grow at a CAGR of 68.7%, while the French market is projected to grow at a CAGR of 67.8%.

North America holds the largest share of the global market, accounting for 48.3% in 2022. This dominance is attributed to factors such as:

- **Increased government funding and investments:** The US government has made significant investments in Digital Twin research and development, driving market growth.
- **Rising need for Digital Twins in drug discovery & development:** The pharmaceutical industry in North America is increasingly adopting Digital Twins to streamline drug development processes and reduce costs.
- **Presence of a large number of Digital Twin companies:** The region is home to many established and emerging companies offering Digital Twin solutions for healthcare, contributing to the market's expansion.

Within North America, the US accounts for the largest share of the market (93.0% in 2022), followed by Canada. The market in Canada is expected to grow at a CAGR of 69.4% during the forecast period.

Asia Pacific is projected to be the fastest-growing region in the Digital Twins in the healthcare market, with a CAGR of 68.5% during the forecast period. This rapid growth is driven by several factors:

- **Growing demand for advanced technologies:** Healthcare systems in the region are rapidly adopting advanced technologies to improve patient care and enhance efficiency.
- **Increasing awareness regarding the benefits of Digital Twins:** Healthcare professionals and stakeholders in the region are becoming increasingly aware of the potential benefits of Digital Twin technology, leading to higher adoption rates.
- **Growing adoption of Digital Twins in emerging economies:** Countries like China and India are investing heavily in digital health initiatives, creating a favourable environment for the growth of Digital Twins in the healthcare market.

China is the largest market for Digital Twins in healthcare in the Asia Pacific region, followed by Japan. The market in China is projected to grow at a CAGR of 69.9%, while the Japanese market is expected to grow at a CAGR of 68.1%.

The **rest of the world**, which includes Latin America, the Middle East, and Africa, accounts for a smaller share of the Digital Twins in the healthcare market compared to North America, Europe, and Asia Pacific. However, the market in this region is expected to grow at a CAGR of 65.6% during the forecast period.

The growth of the market in this region is driven by factors such as:

- **Increasing healthcare expenditure:** Governments in these regions are increasing their healthcare budgets, leading to investments in advanced technologies like Digital Twins.
- **Rising awareness about the benefits of digital health:** Healthcare professionals in these regions are becoming more aware of the potential benefits of digital health solutions, including Digital Twins.
- **Growing investments in healthcare infrastructure:** Several countries in this region are investing in improving their healthcare infrastructure, creating opportunities for the adoption of Digital Twins.

30.1.4 Conclusion of the Landscape Analysis

The DT in the healthcare market is poised for significant growth in the coming years, driven by factors such as increasing investments, expanding applications, and rapid technological advancements. The European market is expected to witness substantial growth, fuelled by EU and Member States' initiatives, rising investments in research & development, and the presence of key market players. The adoption of Digital Twins in healthcare is transforming patient care by enabling personalised treatments, improving surgical outcomes, and optimising healthcare workflows. As the market continues to evolve, even more innovative applications of DT technology can be expected in the future.

30.2 Business models for VHT

30.2.1 Business model strategies and approaches

During the EDITH-CSA project, a wide range of business model strategies were explored through a variety of activities, including modellathon activities, literature study and interviews with key opinion leaders. From this, four broad categories of value proposition emerged in terms of how stakeholders in the VHT ecosystem could capitalise on opportunities presented by the federated infrastructure.

- **Technology development and systems integration:** including OEM (hardware), medical device manufacturers, software application and platform development specialists etc.
- **Data management and security:** a range of value propositions providing a broad range of data management services within the VHT and surrounding ecosystem.
- **R&D and clinical decision support:** stakeholders involved primarily in implementing VHT and DT solutions in clinical environments, often directly engaging patients.
- **Business integration and support:** organisations acting as intermediates to enable the implementation of VHT and DT within different business and operating environments.

Additional categories of value proposition are likely to emerge as the VHT grows and diversifies. In addition, a variety of organisations will be involved across these initial operating segments, ranging from established companies – utilising the VHT within the framework of their existing business model(s) – to SMEs and academic spin-offs, particularly in the first three categories listed above.

In terms of business approaches, three general approaches were recognized:

- *Tech transfer or spinoff businesses (TTSO):* businesses created by academic or research institutions to commercialise innovations. TTSOs leverage university-driven/lab-driven research and IP to bridge the gap between lab discoveries and market-ready solutions for the VHT.
- *Specific tech tools businesses (STT):* R&D-driven private businesses specialising in tools or services that enhance specific aspects of DTs or the VHT. They focus on optimising or supporting specific applications with deep technical expertise.
- *Intermediary businesses (IB):* service providers optimising assistance in various areas (such as compliance, funding, or market strategy). IBs adapt to regional and economic contexts to ensure smooth deployment of VHT technologies.

Box 54: EDITH development – BizMod4DTH portfolio of business models for VHT

EDITH development : a portfolio of potential business models developed for VHT – BizMod4DTH
Source: Deliverable D6.2

EDITH CSA consortium member Pi School created a portfolio (BizMod4DT) of business models deemed to have possibilities to emerge within the VHT ecosystem. BizMod4DT is aimed to encompass a wide array of business cases collectively illustrating how VHT technology can be applied across healthcare. This portfolio explores potential opportunities and assesses the scale and sustainability of the emerging VHT ecosystem highlighting the potential of the VHT initiative to transform patient care, hospital management, personalised medicine, and public health. Each business model was synthesised in brief 'business cards', indicative of a business possibility covering a range of products and services for individual DTs, built using the VHT infrastructure, and for the VHT infrastructure itself.

BizMod4DT was developed first via an iterative, collaborative approach in the form of a Modellathon – a **business model hackathon** – designed to identify innovative, creative and feasible business models tailored to the healthcare sector's growing demand for DT technology. Then, a detailed analysis of the models was

conducted with comparisons from the literature and **interviews of key opinion leaders** to assess real-world applicability and scalability. A **market mapping** of each model of the portfolio was carried out with five KPIs: (i) Alignment with EU healthcare goals, (ii) VHT ecosystem compatibility, (iii) modellathon relevance; (iv) diversity of business sub-categories, and (v) emphasis on VHT-based applications. For each business model, market size estimates were developed using a combination of top-down and bottom-up approaches. These estimates were based on data and forecasts from available data, focusing on the EU healthcare ecosystem. Key metrics, such as cost burden categories, market segmentation, and revenue streams, were derived from publicly available reports, academic literature, and industry benchmarks. When direct EU data were unavailable, assumptions were made using global market shares or data from comparable regions, with adjustments for currency exchange rates, regional healthcare expenditure percentages, and relevant CAGR values.

These analyses resulted in a **selection of 35 business models** that show a wide spectrum of high-level conceptual definitions of opportunities arising from the VHT ecosystem.

The following four tables group the studied business models according to their main value proposition. The last column of each table represents one of the phases at which this type of business is expected to operate: phases 1 (short term, 1-4 years), 2 (middle long term, 5-7 years) or 3 (long term, 8-10 years).

Table 6: Technology Development and Systems Integration businesses and the relative timing to operate: 1 (short term, 1-4 years), 2 (middle long term, 5-7 years) or 3 (long term, 8-10 years).

Technology Development and Systems Integration	
<i>Business model</i>	<i>Phase</i>
Hospital Integration Services specialise in integrating hospitals' IT systems with VHT solutions, focusing on seamless software interface development, continuous system monitoring, and comprehensive compliance support.	3
Application Programming Interface (API) and Software Integration provide APIs and middleware for seamless integration between healthcare systems and virtual twin technologies.	2
Integration of Data Sources and Data provides a platform that enables the integration of genomic, clinical, and diagnostic data into DT models, supporting personalised medicine and genomics research.	2
Patient Engagement and Monitoring provides a comprehensive telemedicine platform that connects patients with healthcare providers remotely using digital twin technology to model patient health data, enabling predictive analytics and personalised treatments.	2
Patient Self-Treatment Tools empower patients to manage their own health and feed data to healthcare providers thanks to self-treatment tools. By using the digital twin technologies, businesses can help provide behaviour change programs that help patients manage their conditions.	2
Precision Medicine Solutions leverage DT created from genomic data, biomarker analysis, and clinical records – to deliver precision medicine solutions using advanced AI-powered predictive software.	2
White-Labelled VHT Applications provide customisable, white-labelled applications built on VHT data, offering healthcare providers ready-to-deploy solutions in areas like chronic disease management, remote patient monitoring, and diagnostics.	3
Medical Device development and data collection provide tools such as wearable monitoring powered by AI algorithms to detect conditions. It feeds real-time, data-rich insights into VHTs, enabling predictive analytics and personalised care.	2
Simulation and Visualisation Tools provide software that visualises complex data and insights for different stakeholders using AI-driven software for automated healthcare data visualisation, enhancing real-time insights for personalised care.	2
Supply Chain Resilience Platform leveraging the VHT technology enhance crisis preparedness by anticipating health needs. The platform analyses health data and early biomarkers to identify risks before symptoms appear, alerting users and public health authorities.	3
Rapid Point-of-Care Diagnostics develop an AI-enhanced, compact blood diagnostic device capable of performing multiple health tests from a single drop of blood, delivering rapid, comprehensive results in 15-30 minutes	2

Table 7: Data Management and Security businesses and the relative timing to operate: 1 (short term, 1-4 years), 2 (middle long term, 5-7 years) or 3 (long term, 8-10 years).

Data Management and Security	
<i>Business model</i>	<i>Phase</i>

Data Brokers/Data Integration Specialist convert unstructured data, including EHR, voice notes and handwritten records, into structured format making it usable in VHT systems.	2
Intelligent Data Management Cloud (IDMC) creates a single source of truth for patient data by integrating EHRs, lab systems, imaging, and wearables with AI-driven capabilities for data integration, standardisation, and governance.	1
Smart Electronic Health Record (EHR) system integrates advanced technologies such as AI and DT capabilities as real-time data processing and consolidates patient information from diverse healthcare sources.	2
Patient Data Ownership Platform empowers patients to own and control their health data within the virtual twin ecosystem focusing on individual data rights.	3
Data Protection and Encryption offers next-generation homomorphic encryption, SMPC, and blockchain-based security to protect sensitive health data enabling secure, VHT-compliant cross-border collaborations.	1

Table 8: R&D and Clinical Decision Support businesses and the relative timing to operate: 1 (short term, 1-4 years), 2 (middle long term, 5-7 years) or 3 (long term, 8-10 years).

R&D and Clinical Decision Support	
Business model	Phase
Epidemiological Modelling utilises VHT technology to aggregate real-time patient data, creating DT for populations enabling precise disease modelling and predictive analysis of epidemic trends.	3
Drug Discovery Acceleration focuses on identifying new drugs leveraging DT technology to simulate drug behaviour across biological pathways, accelerating discovery.	2
Simulation Software for Medical Research provides simulation software that models biological processes for drug research, enabling faster regulatory compliance.	1
Synthetic Trials Using DT leverages AI and DT technology to generate synthetic patient data, simulating responses for virtual clinical trials.	1
AI-Powered Decision-Making Software enhance clinical decision-making by analysing patient-specific virtual twin data predicting treatment outcomes, optimise personalised care, and reduce diagnostic errors.	2
Virtual Reality (VR), Augmented Reality (AR) applications revolutionise surgical planning, training, and real-time navigation delivering personalised, high-precision solutions for complex surgeries, enhancing both surgical outcomes and patient engagement.	2
Neuromorphic computing represents the next frontier in hardware innovation, mimicking brain-like processing for efficient, low-power AI enabling real-time monitoring, rapid diagnostics, and predictive analytics, enhancing patient outcomes and reducing energy costs.	3
Individual Private VHT is a comprehensive digital replica of a person's unique biological profile, securely storing personalised genetic, phenotypic, lifestyle, and biometric data.	2
Burnout Prevention in the Workplace provides a B2B wellness platform leveraging VHT technology and continuous monitoring to assess employee mental health risks, including stress, fatigue, and burnout.	3

Table 9: Business Integration and Support businesses and the relative timing to operate: 1 (short term, 1-4 years), 2 (middle long term, 5-7 years) or 3 (long term, 8-10 years).

Business Integration and Support businesses	
Business model	Phase
Legal and Compliance Advisory Services offer counsel on data privacy, compliance, and malpractice risks associated with VHTs. Legal or ethical oversight committees ensure ethical considerations are addressed in VHT development.	1
Think-tank drives VHT innovation and adoption by connecting researchers, healthcare providers, and industry leaders, fostering knowledge exchange through regular events, workshops, and think-tank activities.	1
Project Management Software centralises all aspects of project execution, regulatory compliance, data sharing, and stakeholder collaboration.	1
Institution & Public Information Service serves as a scientific outreach and education hub, promoting VHT technology across healthcare, academic, and public sectors.	1
Financial and Reimbursement Models support reimbursement strategies based on outcomes between healthcare providers and insurance.	2

Co-creation platform for IoT establishes a B2B co-creation hub where healthcare, IoT, and smart home technology companies partner in the development of integrated health monitoring and predictive wellness solutions.	2
Regulatory Compliance provides a VHT-specific certification and compliance validation program for applications developed on the VHT platform, ensuring adherence to healthcare regulatory standards, data privacy laws, and interoperability requirements.	2
Risk Assessment Services offer specialised risk assessment and advisory services tailored to organisations deploying VHT technologies.	1
Marketplace for model brokers envisions a platform for model brokers that leverages VHT data and simulation technology to offer curated data and code packages.	1
Accelerator & Incubator for Innovation Hubs establish an accelerator and incubator program dedicated to supporting startups, SMEs, and research groups focused on developing VHT applications.	1
Supply Chain Resilience Analysis and Assessment leverage the VHT preparedness by enhancing its capacity to anticipate health needs.	3

30.2.2 Trends & risks

The analysis of the business opportunities identified within the VHT ecosystem was based on two preliminary steps, assessing value proposition trends and the transversal analysis of risks.

The list of **value proposition trends** includes data integration and interoperability, AI-driven healthcare applications, personalisation and precision medicine, regulatory and security solutions, drug discovery and clinical trials, real-time monitoring and predictive care, education, ecosystem building and collaboration, and advanced technological solutions. Such a diversified range of trends highlights the opportunities to create value for investors

The **risks associated** with the 35 business ideas in the selected portfolio include patterns, common challenges, and unique concerns. The risks are categorised into five main areas: regulatory and compliance risks, technical implementation challenges, market adoption pressures, data privacy concerns, and financial and operational risks. In particular, regulatory compliance, technical integration, and cybersecurity are significant concerns that require proactive strategies. Emerging technologies like AR/VR and Internet of Things (IoT) have high potential but face slower adoption and higher financial risks. Business models addressing precision medicine, EHR systems, and drug discovery are better positioned due to strong market validation and alignment with healthcare trends. Mitigation strategies such as partnerships, phased implementation, and regulatory alignment, can help to effectively navigate these risks.

To perform the needed ranking and analysis, the comparable models within each identified business area were analysed, assessing their associated revenue and funding potential. The models were ranked based on the success of their comparable businesses, considering factors like revenue size, funding raised, and market growth. The most promising models appear to be those dealing with precision medicine solutions. Particularly attractive models will be Smart HER systems and Drug Discovery acceleration. All these models show evidence of high scalability and strong alignment with regulatory and market needs. Other promising and emerging opportunities include telemedicine platforms, patient self-treatment tools, synthetic trials, AR&VR applications, co-creation platforms, and risk-assessment services.

30.2.3 A phased approach

When envisioning the timeframe of business model development within the BizMod4DT strategy, the early phases of the VHT implementation will mostly focus on R&D and technology development, along with the ecosystem development. Within this framework, business integration and technology development will occur in early to mid-phases, and clinical decision support will occur in later phases. This timing of course depends on many factors and for every development there will be low hanging fruits and early successes that pave the way.

In the short term (phase 1, 1-4 years), business models will focus on **establishing legally compliant, user-centric foundations** to drive early adoption and innovation. Key models will prioritise compliance with regulatory frameworks to ensure legal robustness, avoiding setbacks and bolstering stakeholder confidence. Emphasis will be placed on creating a user-oriented solution, backed by expert

toolkits – comprising tech transfer officers, legal advisors, and UX/UI specialists – commonly found in accelerators and innovation hubs. This ensures the platform is not only functional but also intuitive and navigable, reducing friction for early adopters like researchers and developers.

Data access and management will be a key business focus, acknowledging its essential role as an incentive. Secure handling of synthetic, anonymised, and pseudonymised data within a well-structured environment will lay the groundwork for long-term VHT credibility and scalability. Parallel to this, an agile and creative ecosystem will emerge, supported by think tanks fostering iterative feedback loops. These loops will enhance the VHT platform, enabling users to share insights, raise ethical concerns, and collaboratively ideate future advancements.

Education and community building will underpin these efforts, ensuring societal engagement and understanding of VHT technologies. Educational initiatives will demystify the VHT's value while addressing ethical dimensions, fostering public trust and proactive participation. Targeting first-hand owners of cutting-edge technology, such as researchers and developers, aligns the VHT platform with the most advanced technological and scientific contributions, reinforcing its position as a future-ready tool.

This phase represents a **low-risk, high-development strategy**: setting secure, compliant foundations, enabling user engagement through a robust and intuitive platform, and catalysing innovation through a collaborative and transparent ecosystem. These efforts will collectively position the VHT as a credible, user-driven solution poised for growth in subsequent phases.

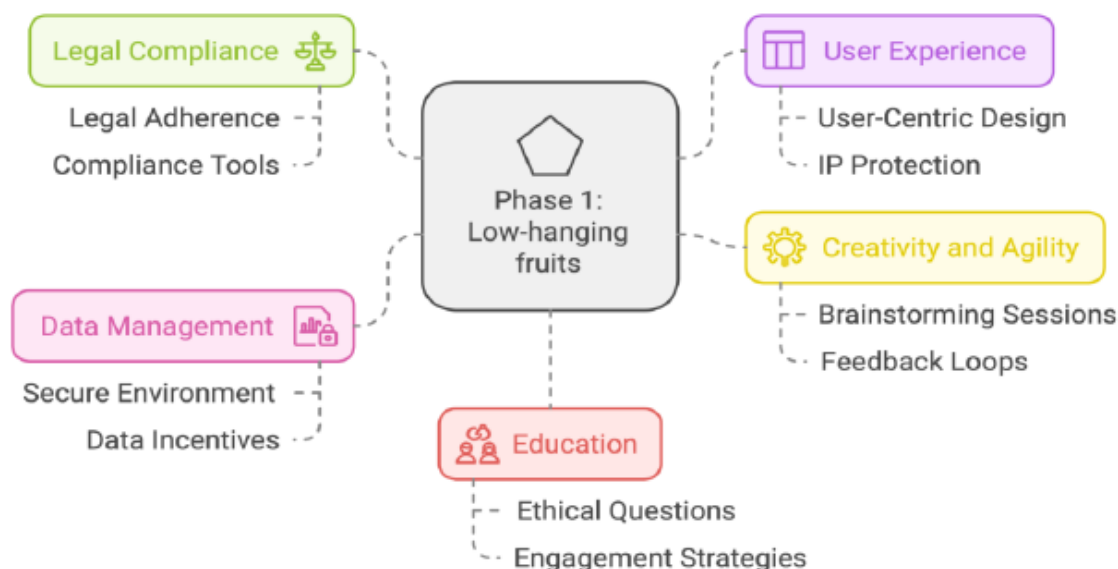


Figure 38: Phase 1 (short term) development of business models.

In the **middle-long term** (phase 2, 5-7 years), business models will focus on scaling the platform, integrating advanced capabilities, and expanding its user base. Building on the secure and compliant foundations of Phase 1, this phase introduces **federated models** to enhance secure data processing, leveraging frameworks like the EOSC. The alignment with the EHDS will enable the secondary use of sensitive data and lay the groundwork for eventual primary data integration. These developments will empower the platform to manage more complex datasets and models, solidifying its credibility and utility.

Business models will capitalise on **pilot implementations to validate real-world applications** of the VHT. By demonstrating clear advantages like better patient outcomes and improved clinical workflows, these pilots will create powerful success stories that draw in new stakeholders, such as hospitals and clinical users. Partnerships with researchers, healthcare institutions, and industry leaders will be central to demonstrating the platform's value and fostering collaboration within the ecosystem.

Tokenization and other **advanced privacy techniques** will form the backbone of secure data management, reinforcing user trust while enabling sophisticated services like automated data analysis.

These enhanced services will cater to diverse user needs, providing actionable insights and increasing the platform's appeal to a broader audience.

Phase 2 business models will also **emphasize maturity in data and model usage**, tackling complex problems that were beyond the scope of Phase 1. By organising collaborative projects and leveraging the expanded capabilities of the VHT, the platform will establish itself as a critical tool in the healthcare and research sectors.

By creating pathways for wider adoption and demonstrating the VHT's value through success stories, Phase 2 will set the stage for even broader integration and sustainability in Phase 3.

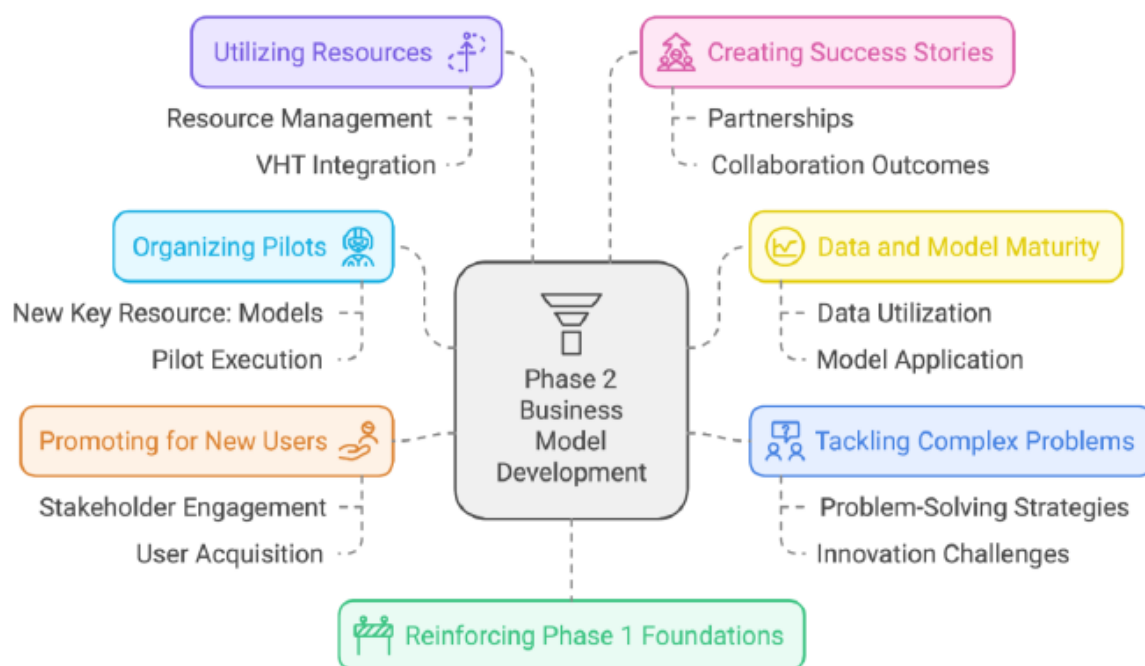


Figure 39: Phase 2 (middle long term) development of business models.

In phase 3 (long term, 8-10 years), business models will shift towards **large-scale deployment and integration into healthcare systems**, unlocking the platform's full potential and enabling high-stakes benefits for stakeholders (Figure 9). This phase will focus on delivering an enterprise-grade solution tailored to hospitals and clinical environments, supported by comprehensive training and technical assistance. By targeting healthcare providers directly, the VHT aims to achieve standardization and widespread acceptance across the EU, breaking down remaining barriers to adoption.

The primary business models in this phase will centre on scalability and sustainability. Enterprises and SMEs will be positioned to **capitalise on the platform's established credibility assessment technologies**, using it to drive significant profits through advanced applications such as personalised medicine, predictive analytics, and population health management. These applications will leverage the VHT's interconnected pipelines of data, models, and simulations to generate actionable insights, meeting the EU healthcare agenda for more efficient and effective health services.

Another critical component will be ongoing system enhancements and **regulatory compliance**, ensuring the VHT adapts to technological advancements and new policy frameworks. These updates will safeguard the platform's long-term viability and reinforce user trust, further embedding the VHT into healthcare workflows.

The **collaborative ecosystem** established in earlier phases will mature into a robust network of stakeholders—including healthcare providers, policymakers, researchers, and technology developers—working together to fully exploit the VHT's capabilities. Mass-scale data analysis will enable the generation of new knowledge and wisdom, driving innovation and continuous improvement across healthcare systems.

In this phase, business models will focus on **maximising the economic and societal impact of the VHT by deploying it at scale**, standardising its use, and fostering profitability for enterprises and

SMEs. This phase marks the culmination of the initiative, where the VHT becomes an indispensable tool in modern healthcare.

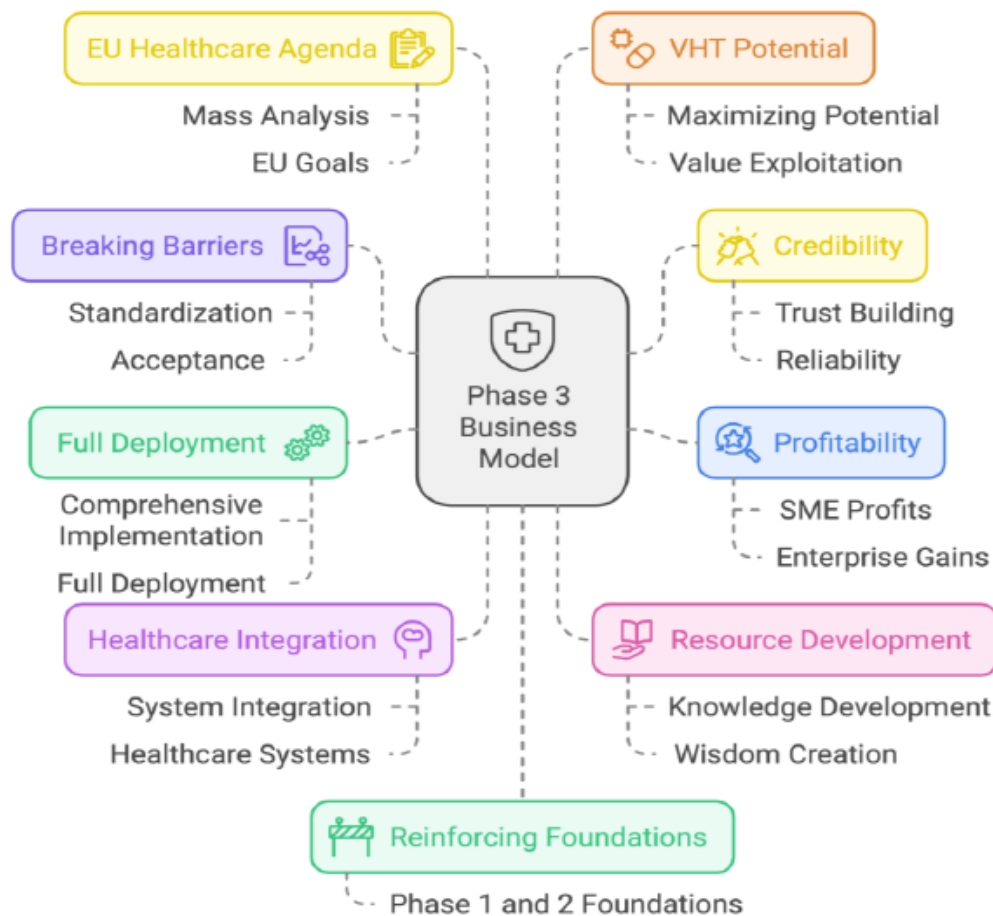


Figure 40: Figure X: Phase 3 (long term) development of business models.

30.3 Digital Twin use cases

Box 55: Use Case - Glycemic control in ICU patients (market)

Use Case : Glycemic control in ICU patients

Status: start-up company, solution in clinical trial

Website: www.insilicare.com

InSiliCare's AI-powered Digital Twin is a clinical decision support solution for ICU patients requiring glycemic control. It provides personalized, safe, and effective management of blood glucose levels and nutrition delivery. Insulin and nutrition treatments are calculated to maximise safety from hypoglycemia, while controlling patient blood glucose levels and optimizing nutrition towards a configurable physician-determined practice of care.

ICU care accounts for about 13-15% of total hospital costs for < 1% of hospitalized patients. In total, 5.7 million patients are admitted every year **in the US**, representing 130 billion USD of care cost. Between 1.7-2.8 million patients (30-50%⁵¹⁶) could thus benefit from our solution in the US only. **In France**, about 1.5 million patients are admitted in the ICU each year (with 5,671 ICU beds in 343 hospitals), where 450,000 - 750,000 would need insulin therapy.

Safe, effective GC significantly improves patient outcomes and reduce length of ICU stay (by about 1 day) whilst simultaneously improving workflow, freeing up resources for other aspects of ICU care, prevent human

⁵¹⁶ <https://doi.org/10.3390/endocrines4030037>

error (50% of medication errors involve insulin, of which 30% are fatal), and provide compliance audit data in case of adverse events^{517,518}.

The ICU care costs per day vary widely based on hospitals and countries. In France and Belgium, the average cost per day in the ICU is 2100-3000€⁵¹⁹. With additional avoided costs from hypoglycemia, human errors, etc., the total cost savings can fairly be estimated around 3000-5000€ per patient. This thus represents well over 1 billion annual care cost savings for a country like France.

Box 56: Use Case - Osteoporotic fracture risk prediction (market)

Use Case: *Bologna Biomechanical Computed Tomography for osteoporotic fracture risk prediction*
Status: *clinical research studies*

Bologna Biomechanical Computed Tomography (BBCT) is a Digital Twin methodology designed to predict the mechanical strength of the femur under critical loading conditions in osteoporotic patients. Quantitative Computed Tomography (QCT) scans of the hip region and patient data inform a subject-specific Finite Element (FE) model able to predict the risk of hip fracture at the time the CT is performed (ARF0).

The main revenue streams for BBCT could come from selling its usage licenses. Assuming the sale of five BBCT licenses per year to five different medical centres, each centre could perform analyses for up to 200 patients annually (based on a total of 1,000 patients per year). Excluding the CT scan cost, which is reimbursed by the National Health System, the cost of a BBCT analysis is estimated at 350 €/patient, resulting in an annual revenue of 350 k€.

Including the CE marking process cost, previously estimated at approximately 146.8 k€/year, the break-even point is reached if each BBCT license is sold for approximately 99.36 k€. If the CE marking cost is excluded, the license cost drops to 70 k€. Further reductions can be achieved by excluding the cost of the Ansys license, assuming it is provided through a partnership or agreement. In this case, the break-even point would be approximately 58 k€.

Storage costs also play a critical role in the cost analysis. Limiting storage to a maximum of 30 GB per patient can reduce the annual storage cost from 60 k€/year to 19.8 k€/year. This brings the minimum BBCT license cost down to 25.5 k€.

Grants, incentives, and funding opportunities could further lower the cost of BBCT analyses. By leveraging its close collaboration with clinicians at the Orthopaedic Institute Rizzoli, the BBCT company/startup could also actively participate in organising, administering, and supporting orthopaedic medicine congresses. These activities provide an additional income stream and present valuable opportunities to reach new customers and expand market outreach.

Box 57: Use Case - Universal Immune System Simulator for Tuberculosis (market)

Use Case : *Universal Immune System Simulator for Tuberculosis*

Status: *augmenting clinical trials*

Website: <https://www.mimesis.srl/>

The Universal Immune System Simulator (UISS) offers a robust, clinically validated platform to address these challenges by simulating and predicting immune responses *in silico*. UISS-TB has demonstrated its utility in optimizing therapeutic strategies for Tuberculosis, evaluating drug regimens, and predicting outcomes.

Globally, TB management incurs substantial healthcare costs and morbidity. UISS-TB's ability to enhance treatment efficacy and reduce trial costs makes it a valuable tool in addressing these challenges. By supporting personalized medicine, UISS-TB can:

- Decrease time-to-market for new therapies.
- Optimize resource allocation in TB treatment.
- Deliver cost savings through improved treatment precision and reduced complications.

To capitalize on its potential, a spin-off from the University of Catania, Mimesis SRL, has been established to commercialize UISS-TB. Mimesis SRL focuses on advancing the platform's development, ensuring accessibility to researchers and clinicians, and integrating it into pharmaceutical pipelines for tuberculosis treatment and vaccine development.

Box 58: Use Case - Epileptogenic zone localisation for surgical planning in epilepsy patients (market)

⁵¹⁷ <https://doi.org/10.1056/NEJMoa011300>

⁵¹⁸ <https://doi.org/10.1016/j.chest.2019.05.016>

⁵¹⁹ <https://doi.org/10.1186/s12913-023-09926-2>

*Use Case : Epileptogenic zone localisation for surgical planning in epilepsy patients***Status:** *in clinical use***Website:** <https://www.cloudsofcare.com/>

Persyst ESI powered by Epilog is a neuroimaging solution that automatically combines scalp EEG data with a patient's MRI to perform Electrical Source Imaging (ESI). It pinpoints the origin of brain activity linked to seizures, helping clinicians accurately localize the epileptogenic zone—critical for surgical planning in epilepsy patients. Clouds of Care NV is a spin-off company of Ghent University, Imec and Antwerp University in Belgium. Epilog is a trademarked brand of the Clouds of Care digital health technology portfolio reflecting the presence in the epilepsy market segment.

Besides providing the ESI-as-a-service solution for surgery planning, the DT solution will be used to further the use of the company's ecosystem platform in clinical development. The solution will be Clouds of Care's in-house digital health technology to be used in clinical trials, together with other best-matched DHTs within the appropriate context of use. This widens the scope and market potential.

1. **Hospitals and Healthcare Providers:** The global market for healthcare analytics and patient monitoring solutions is massive. As of recent reports, the healthcare analytics market is projected to **reach over \$70 billion by 2027**, with growth driven by the increasing adoption of technology in hospitals and healthcare centers.
2. **Pharmaceutical Industry:** Pharmaceutical companies can also benefit from these systems in drug development, especially in areas where real-time patient monitoring and data are essential for clinical trials. The global **clinical trial** market size is expected to grow to over **\$100 billion** by 2030, with more emphasis on incorporating AI and data analysis tools.
3. **Patient Care and Personalization:** The healthcare industry is increasingly focused on personalized medicine, predictive analytics, and improving patient outcomes. As more patients seek advanced care, tools like Persyst ESI powered by Epilog can make a big impact by enabling faster diagnoses and more precise treatments.

*Box 59: Use Case - the Atrial Modelling Toolkit for cardiovascular Digital Twins (market)***Use case : the Atrial Modelling Toolkit for cardiovascular Digital Twins****Status:** *for research purposes***Website:** <https://github.com/pcmlab/atrialmtk>

The Atrial Modelling Toolkit⁵²⁰ (**atrialmtk**) aims to overcome the challenges of constructing cardiac models at scale through the development of a robust, open-source pipeline for bilayer and volumetric meshes for atrial models.

Two key uses are envisaged for atrialmtk: 1. large *in silico* clinical trials of different AF treatment approaches, and 2. personalized AF treatment outcome predictions for an individual. For the latter, in the future, a clinician will enter the demographics of the patient they want to simulate, together with the treatment type they would like to simulate into the simulation platform. This would include different radiofrequency ablation lesion sets and antiarrhythmic drugs, which could be simulated on the atrial Digital Twin at different time points in their disease progression. This model could have different degrees of personalisation including the patient's anatomy, fibrosis tissue distribution, ECG and comorbidities, with coupling of biophysical and data driven approaches to learn from known biophysical elements and large population datasets. The model could predict AF burden for different treatment pathways. Model predictions could be updated using new ECG data from wearables.

⁵²⁰ <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2023.0038>

31 Towards a VHT marketplace

In addition to the business modelling approach and analysis reported in the previous section, this section takes a step aside to explore a **hypothetical vision for a centralised and federated VHT marketplace**, offering potential pathways for enhanced collaboration and resource sharing among users. It provides insight into establishing a marketplace within the VHT. It outlines the different adoption phases, the technology utilised, and pricing strategies for resources available on the VHT platform.

31.1 Adoption Phases

Considering the intricate nature of the involved community of practice and their diverse expectations regarding the value propositions of the VHT, the VHT ecosystem will be an evolving and flexible entity. We can delineate three distinct phases that build upon the initial realisation of a Distributed Ledger Technology (DLT) system within the VHT infrastructure (depicted in figure below). This DLT system would enable the permanent tracing of all types of assets exchanged on the platform while also tracking their origins. Additionally, it would ensure the discoverability, accessibility, semantic interoperability, and reusability of all activated resources.

- During the initial stages, the DLT infrastructure would exclusively accommodate pre-competitive transactions. Incentives would be based on a quality scoring system, referred to as the “**Honour ledger**”.
- Moving to the second stage, both pre-competitive and competitive transactions would coexist, facilitated by the issuance of digital tokens by the DLT infrastructure. These tokens would possess no direct monetary value but would serve as a framework upon which symbolic prices can emerge through the dynamics of supply and demand for all traded assets, including DLT services. This stage is referred to as the “**Token ledger**”.
- Finally, in the third and final stage, the ecosystem would mature and specialise. While some entities would continue to engage primarily in pre-competitive transactions, an increasing number of participants would shift their focus towards competitive business-to-business exchanges, where prices would be set in Euros rather than tokens. This phase is denoted as the “**Money marketplace**”.

These phases are discussed in more detail below.

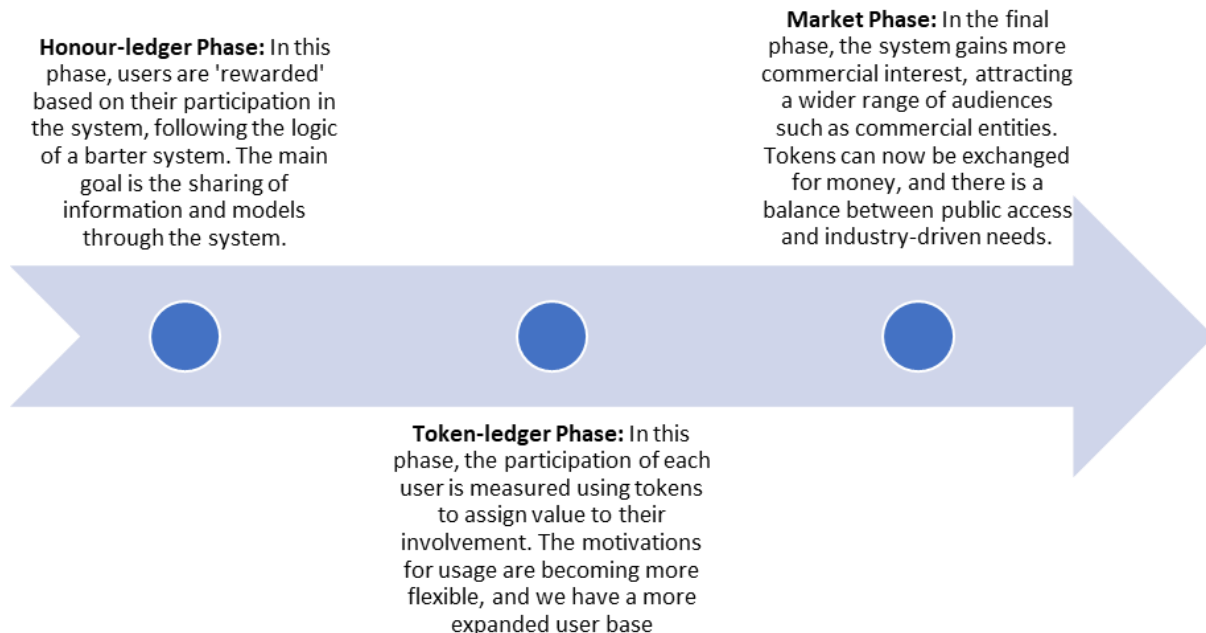


Figure 41: The three phases of the VHT marketplace adoption.

31.1.1 Honour-ledger phase

The currency used in this phase is only a “quality” score based on the reciprocal assignment of quality scoring by all the partners involved. The ecosystem de facto operates as a barter mechanism facilitated by the operation of the DLT infrastructure. By tracing all transactions and quality outcomes, the ledger will be entirely funded by public resources.

The focus in this phase will be on **interoperability**, *e.g.*, how the **participants can barter data and models** with the smallest possible effort. The DLT infrastructure should make it very easy to share a dataset or a model in ways that are compliant with the FAIR (*Findable, Accessible, Interoperable, and Reusable*) principles.

The main motivations for sharing data and models will be funders forcing them to do so. Still, also because it allows them to barter access to their resources with others they can use, and because the use of resources by others will be tracked, possibly with mechanisms that translate them into citation-based reward systems and quality scoring.

The main motivation for reuse will be the simplification of developing and validating new models. This should be particularly true in developing multi-scale and multi-system models, where the work of one on a sub-system can be entirely re-used.

In this phase, the **metric of success is how much data and how many models are shared**. This translates into another sub-metric: the pain-gain ratio for contributors. The easier it is to share resources, and the more rewarding it is, the better it is.

The **governance will require a reverse “T” model**: the day-by-day operations will be ensured by a single organisation or a small consortium that is paid to do it and take technical decisions in an autocratic way. Any other decision on the infrastructure is made using a direct democracy model, where all contributors can participate in the decisional processes through an assembly process.

31.1.2 Token-ledger phase

Eventually, the ecosystem will see the **flourishing of competitive transactions**, aside from the pre-competitive ones. The value of both will be expressed in **digital tokens** issued by the legal entity governing the DLT infrastructure. The “cashing-in” of tokens, by operators contributing resources to the ledger exchange system and the “paying-out” of tokens, by operators purchasing assets offered by others, will allow the development of an increasing nexus of token prices for all transactions taking place through the DLT, which will also begin to charge a token-fee for its services.

The ledger infrastructure will remain a substantial public resource, but its governance will require a more articulated democracy with representation.

The VHT ecosystem will use its **growing token economy** to explore and experiment with how it can become **progressively self-sustaining**. In a more advanced stage of development, the DLT infrastructure will possibly also engage in analysing how incentives linked to automated assignment and distribution of value can be determined by ML mechanisms valuing different attributes or even through Shapley Values mathematical methods (inspired by Nobel Laureate Lloyd Shapley) determining the only distribution satisfying a collection of properties within a coalition game.

In its basics, the token-ledger phase will be characterised by **token-based exchanges replacing barters**, and tokens will be issued to whoever contributes resources to the ledger. However, it will also be possible to purchase tokens in exchange for money: this will apply mainly to external entities not having contributed resources to the ledger but wishing to use the DLT facilities. Further tokens will be gained anytime shared resources will be used, while everybody will pay with the tokens they have accumulated for being allowed to use somebody else’s resources.

The focus in developing the infrastructure will be the **quality of service for its many users**. Systems must scale to extensive collections and handle a truly distributed system based on various hardware and software providers.

The **main motivations for sharing will remain the same as in the previous phase**. However, the move from an honour ledger to a token ledger will render the incentives for what one can get in return for his/her shared resources much more fine-grained and flexible.

The **main motivation for reuse will change** by the extent to which the VHT ecosystem will now be influenced by the development strategies of research groups and companies that have purchased tokens and are willing to use the DLT.

This, of course, implies guarantees of persistence for the infrastructure. In this phase, the **metric of success is how important the VHT becomes** in the development strategies of public and private developers.

Because of the **need to ensure long-term sustainability**, the VHT infrastructure will have to be **run by a legal entity**, possibly organised as a non-governmental organisation (*e.g.*, a foundation), a European Digital Infrastructure Consortium (EDIC), or as a joint undertaking between the EC and the

major European industrial players, like EuroHPC (European High-Performance Computing), IHI (Innovative Health Initiative), etc. This organisation will need to ensure the existence of a public segment of the VHT for the not-for-profit researchers, where most interoperability technicalities standards are tested and standardised. However, it will also have to favour the creation of fully commercial segments of the VHT, which are certified for interoperability by the leading legal entity. Beyond that, they operate entirely privately, pursuing sustainable business-to-business models.

31.1.3 Market phase

Eventually, the number of transactions triggered by research groups and commercial companies having **monetised their access to the DLT-operated ecosystem** will possibly end up changing the nature of the latter, attracting a growing number of subjects out of the ledger and into a **fully-fledged money marketplace**. Academics will remain the artisans exploring the borders of the VHT territory, while the merchants and entrepreneurs increasingly tend to privatise the ecosystem. Will the ecosystem thus dissolve into a mature industrial sector, or will the DLT infrastructure maintain sufficient resilience and attractiveness because of the advanced qualities of the services it will provide?

In any case, one should expect there to be **a public VHT** for academic research and early pre-competitive developments, supported by the EC like any other research infrastructure, **various VHTs for not-for-profit activities**, supported by various charitable mechanisms, and **several commercial VHT infrastructures** that provide Business-to-Business services to an ever-growing industry of *insilico* medicine.

31.1.4 Stakeholder Engagement Across Ecosystem Phases

Different **stakeholders may be more likely to contribute to or join the ecosystem at different stages**. For example, commercial entities and established industries in healthcare may be less certain until key elements of the VHT are in place to ensure data security, integration, support, as well as resource valuation. The **Honour-Ledger Phase will attract academics, non-profit research institutions, and publicly funded projects**. The goal of this phase is to use data from different entities in the context of a barter system, with funding for this phase relying on public sources. The focus of this phase is to attract participants based on the transparency of the resource-sharing system, which is grounded in the principles of open science. The **Token-Ledger Phase will attract additional private research groups and industry stakeholders**. It will be based on a token economy, where each entity earns tokens based on its contribution to the overall platform. The motivation here is the financial reward tied to each entity's participation in the marketplace, which will appeal to those looking for access to the system's resources. This phase is ideal for the creation of public-private partnerships. The **final phase will attract commercial and industrial entities, as well as entrepreneurs**, and will focus on a monetised ecosystem where different entities can buy and sell resources. This phase will be based on the exchange of tokens for financial compensation and will attract entities looking for access to sustainable infrastructure in order to gain financial rewards and competitive advantages.

Table 10: Stakeholder Engagement Across Ecosystem Phases.

Phase	Payment System	Entities Attracted	Motivation
Honour-Ledger	Quality Score-Based System	<ul style="list-style-type: none"> Academics Non-Profit Research Institutions Publicly Funded Projects 	Complying with resource-sharing requirements, earning recognition, advancing open science
Token-Ledger	Token-Based Economy	<ul style="list-style-type: none"> Public Research Groups Private Research Groups Industry Stakeholders 	Earning exchangeable tokens, accessing scalable systems, building collaborations
Market	Monetised Marketplace with Direct Financial Transactions	<ul style="list-style-type: none"> Commercial Entities Established Industries Entrepreneurs 	Financial gains, competitive advantage, accessing scalable system

31.2 Distributed Ledger

In crafting a marketplace tailored for VHT, the VHT platform must address various key requirements to ensure its efficacy and reliability. Firstly, the platform demands **enhanced security measures to safeguard sensitive healthcare data and transactions**. Features such as immutable transactions facilitated by blockchain technology can significantly mitigate the risk of fraud by permanently recording all interactions. Moreover, **transparency and trust** are fundamental pillars that a VHT marketplace must uphold, necessitating a system that enables transparent records and traceability of transactions. Finally, **decentralisation** is important to promote autonomy among users and eliminate the dependence on a central authority, thereby fostering a more resilient and efficient marketplace ecosystem.

Blockchain emerges as the optimal solution for a VHT marketplace due to its unique capabilities and benefits. Firstly, blockchain's immutable ledger ensures that all transactions are securely recorded, preventing tampering or alteration of critical data. Additionally, blockchain's transparent and decentralised nature builds trust among participants by allowing every transaction to be publicly recorded and verified. The distributed network architecture eliminates the need for a central authority, reducing the risk of a single point of failure and promoting peer-to-peer interactions. Furthermore, blockchain's traceability feature enables comprehensive tracking of product transactions, enhancing authenticity and combating counterfeiting within the healthcare marketplace. Overall, blockchain technology offers a secure, transparent, and decentralised framework ideally suited to meet the complex demands of a VHT market. Sensitive information, such as personal health data, remains protected outside the blockchain environment (*e.g.*, securely held by the original data controllers). By isolating sensitive information from the blockchain and leveraging its decentralised architecture for transaction verification, one can foster transparency and trust in transaction records without compromising confidentiality.

Blockchain, a type of DLT, is a sequence of blocks of data records stored in a distributed, transparent, and immutable manner. Originally developed for Bitcoin, blockchain uses unbreakable cryptographic schemes to create secure architectures that enforce privacy while decentralising the decision-making process⁵²¹. Essentially, applications built using blockchain realise a trust-less system that incentivises each participant to make the right decision and not act with malicious intent.

Various ways have been developed to **achieve decentralisation and trust in blockchain applications**, called **consensus mechanisms**. In short, consensus mechanisms are protocols that ensure nodes are synchronised and agree on the validity of transactions that get appended to the blockchain. The first consensus mechanism, used in Bitcoin and later in many other projects, is known as Proof of Work (PoW). Other mechanisms have also been developed, including Proof of Stake, Delegated Proof of Stake, etc.

31.2.1 Blockchain networks

In the last decade, multiple big tech companies have been interested in blockchain technology. This has generated large amounts of capital in research for new use cases and inventions in this sector. Subsequently, this exploration has given birth to several types of blockchain networks varying depending on the use case's needs. For example, projects using sensitive data or decentralising the supply chain management cannot be open to the public, nor can participants be anonymous. The two main types of blockchain networks are public or permissionless and private or permissioned.

Public blockchain network architectures constitute most of the current deployed projects. In general, their key aspects consist mainly of these features: (i) Participants are anonymous; (ii) Anyone can participate without a need for authorisation or verification; (iii) There must be an incentive for the participant to utilise consensus. The most common public blockchains are Bitcoin and Ethereum, which use native cryptocurrencies as an incentive integrated into their consensus protocols.

Private or permissioned networks offer blockchain solutions to use cases where a public blockchain would never be sustained. Important core features of private blockchains include: (i) participants need to be verified and approved to participate on the network; (ii) a member's identity is always known,

⁵²¹ Nakamoto, Satoshi, and A. Bitcoin. "A peer-to-peer electronic cash system." Bitcoin. <https://bitcoin.org/bitcoin.pdf>

acting as a malicious act blocker; (iii) there may or may not be a native currency, depending on the implementation.

Private networks take the core features of the blockchain, meaning data immutability, transparency, and consensus, and apply them to a use case where privacy and high throughput in terms of transactions are more important. This is achieved by implementing alternative consensus mechanisms that are much faster compared to others used by permissionless networks such as PoW, given the fact that it is much more unlikely for an identified member to act maliciously, and even if they do, the issue would be recoverable in general terms.

31.2.2 Smart Contracts

A smart contract is a piece of **self-executing code representing an agreement between the two sides of a transaction**. Firstly, introduced by Nick Szabo⁵²², smart contracts function as a digitally defined protocol that guarantees the outcome will be what it has promised. The substantial evolution of smart contracts came with Vitalik Buterin, founder of Ethereum⁵²³, who made the establishment of smart contracts possible on the chain.

This technology can be used to remove the middleman from many services while maintaining trust in the overall process. Also, every transaction is recorded on the blockchain immutably, making them traceable and irreversible. The most common use cases include voting systems, auctions, asset exchanges and others.

In the context of the VHT marketplace, different types of smart contracts are envisioned, each providing distinct rights and facilitating different types of user transactions. These contracts will offer varying levels of access to the data and resources within the VHT marketplace. Smart contracts will be tailored to meet the project needs at each phase of the marketplace.

Smart contracts that define the ownership of assets: a dataset, model, analytics or simulation service, or a tool can be designated as belonging to a specific entity within the VHT ecosystem. Depending on the type of asset, its transferability can be restricted. For example, a dataset containing sensitive personal information about patients should not be transferable.

Smart contracts that govern the licensing of assets: these contracts will allow an entity that owns an asset to grant usage rights to another entity within the VHT ecosystem. For instance, an entity may be granted permission to use a dataset to train a model. To ensure privacy, sensitive data will only be accessed and processed within the VHT platform. This ensures that an entity can train a predefined model without gaining full access to the underlying data. Federated learning techniques will enable model training without providing direct access to sensitive information.

Different usage scenarios will apply based on the type of asset. For example, a research organisation could use a dataset to create new models or to perform statistical analyses. A doctor might use machine learning or computational models for support in drawing conclusions about a patient's condition. A researcher could leverage tools for analysing datasets and running simulations on DT models. An educational organisation might utilise training and educational materials provided within the VHT platform. Each usage scenario will be governed by specific restrictions regarding the asset's use.

31.3 Equitable Resource Pricing

In a digital marketplace, fair allocation of revenues is critical to fostering trust, collaboration, and sustainable growth among participants. Equitable resource pricing ensures that each contributor's efforts are recognised and rewarded accordingly, motivating active engagement and cultivating a sense of ownership. A transparent revenue allocation marketplace facilitated by clear rules and mechanisms builds trust and confidence among participants, reducing potential disputes and conflicts. Additionally, efficiency in resource allocation, guided by insights, optimises the marketplace's overall performance, maximising value creation and enhancing the user experience. By prioritising these properties, a digital marketplace can cultivate a trusting ecosystem where stakeholders can innovate.

In the VHT ecosystem, “**resource holders**” such as hospitals, upload segments of their medical records to the VHT platform, while “**resource users**” such as clinicians or industries engaged in personalised

⁵²² Szabo, Nick. "Formalizing and securing relationships on public networks." First Monday (1997).

⁵²³ Buterin, Vitalik. "A next-generation smart contract and decentralized application platform." white paper 3.37 (2014): 2-1.

medicine, pay a designated fee to access patient data. One of the hurdles in such data markets lies in determining the fair allocation of payments from analysts back to the contributing entities.

For example, let us assume a VHT model is trained through a federated learning process, where multiple data providers contribute their datasets during training while the data privacy of each contributed dataset remains uncompromised. When the resulting model acquires commercial value, an important question arises: *“How can we fairly determine and assign value to each provider’s contribution during the training process in a transparent and meaningful way?”*.

31.3.1 A game-theoretic framework

An approach often used to address the aforementioned challenge is through a game-theoretic framework, wherein the data contributors are depicted as the players in a coalition game, and the data’s utility from any contributor subset is defined by a utility function.

The Shapley Values⁵²⁴ serve as a classic method in cooperative game theory for **distributing the total gains generated by the coalition of all players**. The widespread adoption of Shapley Values stems from its unique profit allocation scheme, which adheres to a set of properties with practical interpretations, including fairness, rationality, and decentralisation⁵²⁵.

In the context of VHT marketplaces, Shapley Values offer a robust framework for ensuring fair revenue distribution and transparency. Shapley Values can **guide the design of incentive structures** that motivate participants to contribute positively to the marketplace. For instance, in data-sharing initiatives, entities who contribute valuable data can be rewarded based on their Shapley Values, encouraging continued participation and collaboration and promoting trust and cooperation among stakeholders.

The challenge in adopting the Shapley Values lies in their **computational cost**. Evaluating the exact Shapley Values involves computing the marginal utility of every user to every coalition, which is exponential in the size of the number of contributors. Instead of computing the actual Shapley Value, there are various efficient techniques to compute approximations of the true SV. This is a growing research trend, with significant ongoing work being directed toward this area⁵²⁶.

31.4 A Use Case & the Marketplace Architecture

This section analyses the parameters of the VHT Marketplace by providing a **simple use-case scenario** and discussing the system architecture. In such a marketplace designed specifically for the VHT, blockchain technology plays a crucial role in ensuring secure, transparent, and efficient transactions. Let us walk through an example where a healthcare provider seeks to purchase access to a VHT dataset for research and development purposes.

- First, a **data provider**, such as a hospital or research institution, uploads the VHT dataset within the VHT platform. Each dataset is assigned a unique cryptographic hash to ensure integrity and traceability. Then, the data provider lists the dataset on the marketplace, specifying the price, usage terms, and detailed description of the dataset.
- Next, a **healthcare provider** looking for specific VHT datasets to enhance their research browses the marketplace and selects the desired VHT dataset. After agreeing to the terms of use, the healthcare provider initiates the purchase by sending a transaction request to the blockchain. This request triggers the creation of a **smart contract** on the blockchain, which outlines the agreed price, terms of access, and payment conditions and includes the unique hash of the dataset.
- The healthcare provider then **transfers the payment**, either in cryptocurrency or digital tokens, to the smart contract. The smart contract holds the payment in escrow until the transaction is verified and complete. Upon successful verification, the smart contract grants the healthcare provider access permissions to the dataset. The healthcare provider receives the necessary keys to access the VHT dataset

⁵²⁴ Shapley, Lloyd S. "A value for n-person games." (1953): 307-317.

⁵²⁵ Ghorbani and Zou were the first to introduce SVs as a measure for the Equitable valuation of data for machine learning: A. Ghorbani and J. Zou, "Data shapley: Equitable valuation of data for machine learning," pp. 2242–2251, 2019.

⁵²⁶ Maleki, Sasan. Addressing the computational issues of the Shapley value with applications in the smart grid. Diss. University of Southampton, 2015; Jia, Ruoxi, et al. "Towards efficient data valuation based on the Shapley value." The 22nd International Conference on Artificial Intelligence and Statistics. PMLR, 2019; Chen, Hugh, et al. "Algorithms to estimate Shapley value feature attributions." Nature Machine Intelligence 5.6 (2023): 590-601; Wang, Jiachen T., and Ruoxi Jia. "Data Banzhaf: A robust data valuation framework for machine learning." International Conference on Artificial Intelligence and Statistics. PMLR, 2023.

within the VHT platform. With access granted, the healthcare provider confirms receipt and usability of the data. The smart contract then releases the payment to the data provider, completing the **transaction**. Each step of this transaction is immutably recorded on the blockchain, ensuring a transparent and traceable process. The healthcare provider does not have immediate access to the actual VHT dataset; instead, they receive access to use the dataset. This means they can develop models, simulate patient health conditions, enhance personalised treatment plans, and perform statistical analyses without accessing the raw data, addressing privacy complications. By maintaining data privacy, the healthcare provider can leverage the dataset's insights to drive innovative research while adhering to stringent data protection regulations.

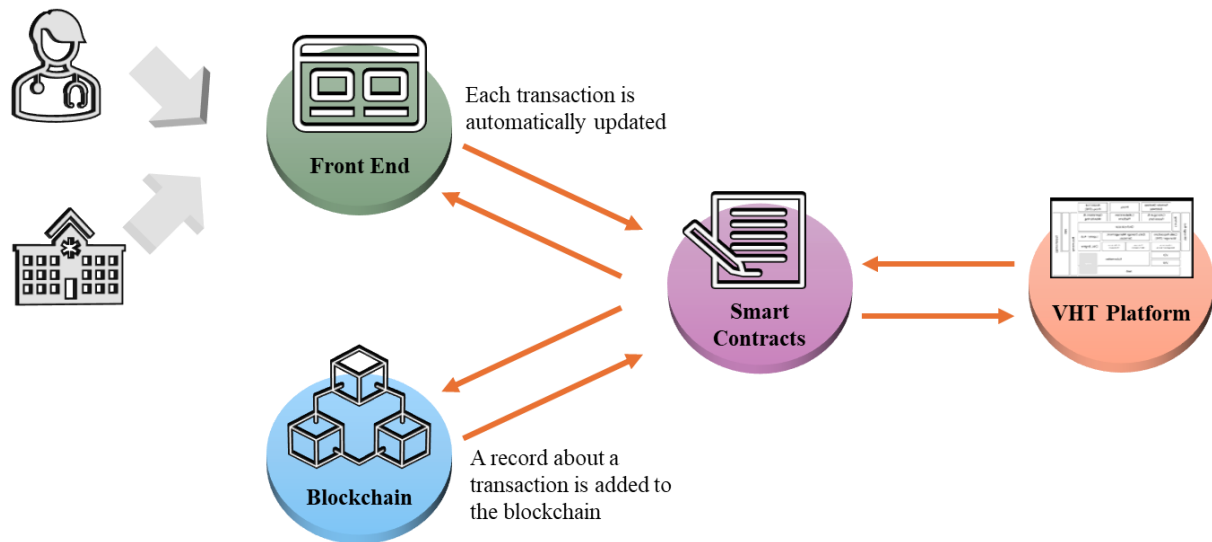


Figure 42: VHT Marketplace Use Case.

In the **architecture of a decentralised application** for a digital healthcare marketplace on top of the VHT platform, the internal structure is presented in the next figure. The decentralised marketplace application follows existing best practices presented in the bibliography⁵²⁷.

The Marketplace ensures data integrity and facilitates interactions with datasets, models, and algorithms, all while maintaining strict access controls to the underlying information⁵²⁸. Each uploaded item's hash is stored on the blockchain to ensure data integrity and verification, while the VHT platform manages the actual data. This approach avoids the inefficiency of directly storing large files on the blockchain and optimises the VHT platform's specialised capabilities in healthcare data management. Furthermore, it enables the development of tools, models, and statistical predictions using sensitive personal data while preventing unauthorised reuse or exploitation of the data by third parties once it leaves the VHT platform. This ensures sensitive personal data from patients is safeguarded effectively.

⁵²⁷ Ranganathan, Vishnu Prasad, et al. "A decentralized marketplace application on the ethereum blockchain." 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC). IEEE, 2018.

⁵²⁸ Raouzaoui, Amaryllis, et al. "Deliverable 5.1: Simulation platform Report." EDITH: Building the European Virtual Human Twin.

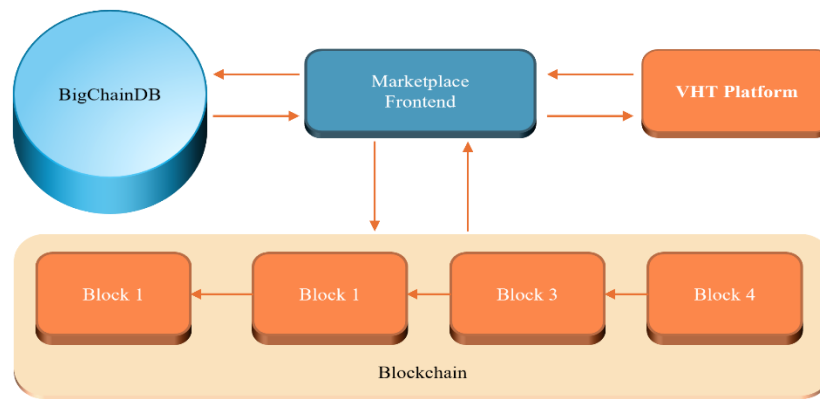


Figure 43: VHT Marketplace Architecture.

This blockchain-based transaction model offers **several benefits specifically tailored for VHT**. **Transparency** is enhanced as every transaction step is recorded on the blockchain. Security is ensured through smart contracts that manage payments and data access securely. **Data integrity** is maintained using hash-keys for files and blockchain, guaranteeing the data remains unaltered and accurately traceable. **Decentralisation** reduces the risk of data breaches and enhances trust among participants. Lastly, the efficiency of the process is increased by automating transactions through smart contracts, reducing the need for intermediaries and manual intervention.

By leveraging blockchain technology, the digital healthcare marketplace ensures that VHT datasets are exchanged in a secure, transparent, and efficient manner. The use of oracles for data verification and IPFS (InterPlanetary File System) for decentralised storage further enhances the reliability and integrity of these data exchanges, fostering a trusted environment for healthcare innovation and collaboration. This structure allows for advanced simulations and personalised care models, driving forward the potential of DT.

32 European public infrastructure

There are numerous Research Infrastructures (RIs) across various fields that are closely related to the VHT. Some focus on specific components or features of the VHT, while others are more general, providing critical support to stakeholders involved in VHT activities. **As of now, a dedicated VHT infrastructure does not yet exist.**

EDITH-CSA's research efforts have implied an extensive examination of a wide range of platforms, tools and repositories at both European and global levels. The related ecosystem has fostered various forms of collaboration with multiple initiatives, usually characterised by a strong emphasis on the lawful and ethical sharing and access of data and services across different scientific disciplines, facilitated by cogent legal frameworks. Summarised below is an analysis of the state of the art of European research infrastructures, followed by the identification of initiatives in Europe and at member state level being relevant for the VHT. This chapter contributes to defining the requirements of a VHT dedicated infrastructure.

32.1 State of the art European infrastructures

In previous parts of this roadmap, a wide range of European (research) infrastructures were already discussed as specific examples or partners for specific tasks. These include, amongst others, EBRAINS, ELIXIR, EATRIS, BBMRI-ERIC, EOSC, OpenAIRE, Gaia-X, and they will not be discussed again. Several other relevant initiatives are discussed in this section.

The **EUDAT Collaborative Data Infrastructure (EUDAT CDI)**⁵²⁹ is one of the largest infrastructures of integrated data services and resources supporting research in Europe. EUDAT envisions sharing and preserving data across borders and disciplines by enabling data stewardship within and between European research communities through a Collaborative Data Infrastructure (CDI), a common model and service infrastructure for managing data spanning all European research data centres and community data repositories. EUDAT offers heterogeneous research data management services and storage resources, supporting multiple research communities as well as individuals, through a geographically distributed, resilient network and data is stored alongside supercomputers.

The **European Federation for Cancer Images project (EUCAIM)**⁵³⁰ aims to deploy a pan-European digital federated infrastructure of FAIR cancer-related de-identified images from Real-World. EUCAIM infrastructure will preserve the data sovereignty of providers and provide a platform for developing and benchmarking AI tools. EUCAIM focuses on the field of cancer, includes an Atlas of Cancer Images, pathology, molecular and laboratory data and will address the fragmentation and cluster the existing cancer image repositories by exploiting the AI4HI initiative⁵³¹, European Research infrastructures, regional and national repositories.

The **Genomic Data Infrastructure**⁵³² (GDI) project enables access to genomic and related phenotypic and clinical data across Europe. It does so by establishing a federated, sustainable and secure infrastructure to access the data. It builds on the outputs of the Beyond 1 Million Genomes (B1MG) project and is realising the ambition of the 1+Million Genomes (1+MG) initiative. A global grass-roots-led open initiative in this same space, is the **Human Cell Atlas**⁵³³ (HCA), joined by over 3000 scientists across the globe and supported by many funders from all around the world. The HCA Data Portal stores and provides single-cell data contributed by labs around the world. The cloud-based platform houses community generated, multi-omics, open and managed access data. The Data Explorer interface acts as the heart of the Portal, to help find and explore HCA datasets.

The **Research Data Alliance**⁵³⁴ (RDA) was launched as a community-driven initiative in 2013 by the European Commission, the United States Government's National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation with the goal of building the social and technical infrastructure to enable open sharing and re-use of data.

⁵²⁹ <https://www.eudat.eu/>

⁵³⁰ <https://cancerimage.eu/>

⁵³¹ Kondylakis H *et al.* Eur Radiol Exp (2022); 6: 29. <https://www.doi.org/10.1186/s41747-022-00281-1>

⁵³² <https://gdi.onemilliongenomes.eu/>

⁵³³ <https://www.humancellatlas.org/learn-more/#event-launch-of-the-human-cell-atlas>

⁵³⁴ <https://www.rd-alliance.org/about-rda>

RDA has several working groups that are highly relevant to the VHT, including the Immune Digital Twin WG⁵³⁵ and the blockchain applications in health WG⁵³⁶.

As can be appreciated from the discussions of the existing initiatives above and elsewhere in the roadmap, there is a rich diversity in terms of infrastructures, initiatives and the services they offer. It is paramount that the VHT, as prospective player in this busy playing field, does not aim to duplicate existing efforts and offers. Rather, it should develop its own unique selling proposition for a specific set of activities, while collaborating with other initiatives for existing services.

32.2 ERICs and EDICs

32.2.1 European Research Infrastructure Consortium

In the context of European infrastructures for scientific research, over the recent years there has been a **trend towards federation of resources** provided by multiple academic institutions and their partners, across national borders. Today, there is a large number of Research Infrastructures (RIs) and the number is growing fast due to the role of science as a driver of innovation and economic growth, the ever-increasing importance of information technology in science, and the benefits that accrue from greater collaboration and scale. The European Commission has foreseen the instrument of an ERIC⁵³⁷ (**European Research Infrastructure Consortium**) as a mechanism to host these RIs.

In the 2021 roadmap⁵³⁸ of the European Strategy Forum on Research Infrastructures (ESFRI), there are 16 RIs in the Health and Food category and 4 in Data, Computing and Digital Research Infrastructures (see figure below), all 20 of them of distributed type. The **European Life Science Research Infrastructures** (LS RIs)⁵³⁹ focus on providing researchers the ability to use highly complex technologies and access services and resources to world-class facilities, samples, instruments and data focusing on Life Sciences. These LS RIs support cutting edge science and offer access to resources and services to all (European) scientists from academia to industry. A Catalogue of Services⁵⁴⁰ for LS RIs has been developed by the CORBEL project⁵⁴¹, providing users with the ability to search for a plethora of needs: from Data and Databases to Technologies & Facilities and Models & Tools. The figure below also shows the interactions between the existing LS RIs in past and current projects.

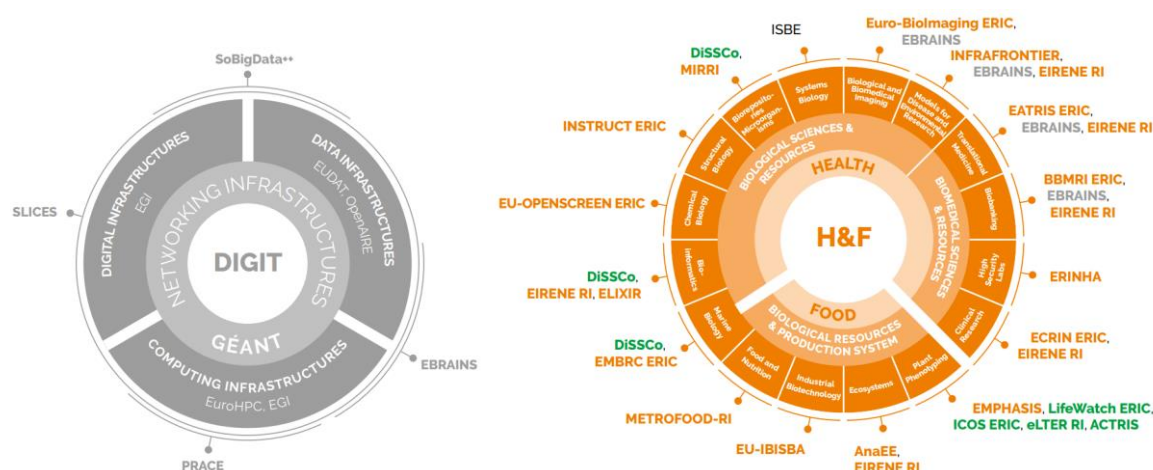


Figure 44: Schematic representation of the 4 Data, Computing and Digital Research Infrastructures (left, grey) and the 16 RIs in the Health and Food category (right, orange) as depicted in the 2021 roadmap of European Strategy Forum on Research Infrastructures.

⁵³⁵ <https://archive.rd-alliance.org/groups/building-immune-digital-twins-wg>

⁵³⁶ <https://archive.rd-alliance.org/groups/blockchain-applications-health-wg>

⁵³⁷ https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/european-research-infrastructures/eric_en

⁵³⁸ <https://www.esfri.eu/esfri-roadmap-2021>

⁵³⁹ <https://lifescience-ri.eu/ec-funded-projects.html#c1940>

⁵⁴⁰ <https://lifescience-ri.eu/catalogue-of-services.html#ui-id-1>

⁵⁴¹ <https://www.corbel-project.eu/home.html>

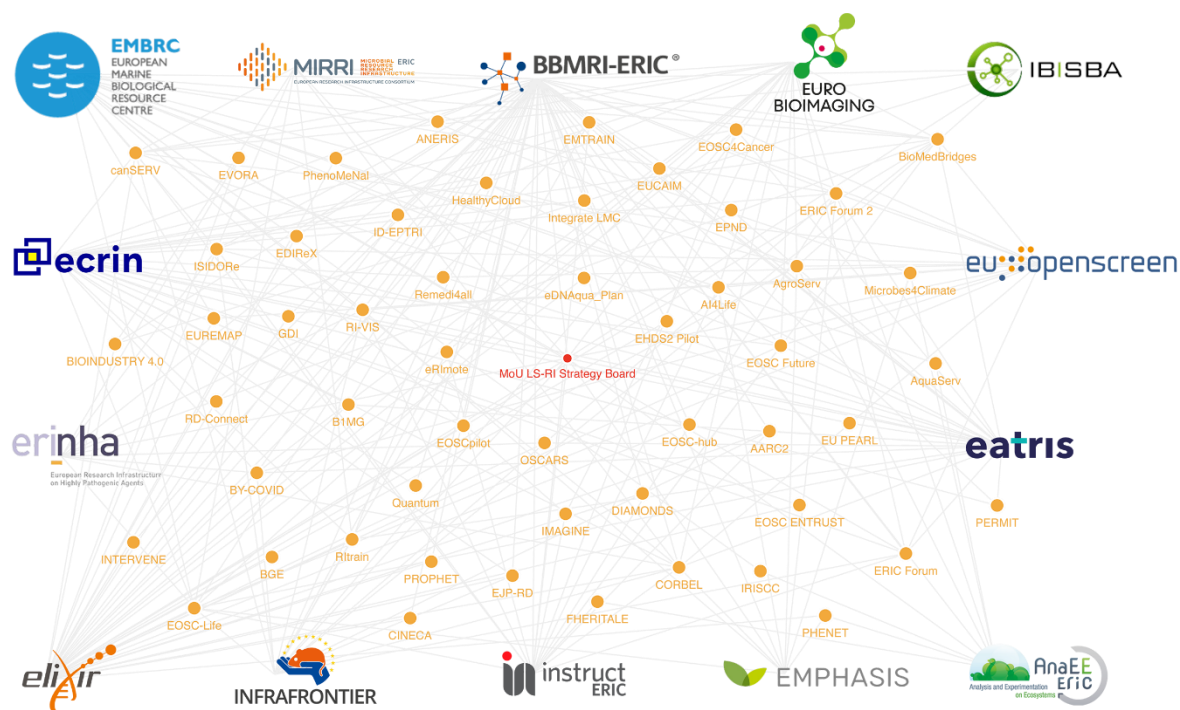


Figure 45: Overview of the Life Sciences Research Infrastructures and the various projects through which they collaborate to increase impact.

32.2.2 European Digital Infrastructure Consortium

In the context of the Digital Decade Policy Programme, the European Commission created another instrument to speed up and simplify the setup and implementation of multi-country projects, called EDICs⁵⁴² (**European Digital Infrastructure Consortium**). Each EDIC is a legal person established by a decision of the European Commission based on an application submitted by at least three Member States. The **founding Member States** define the EDIC's governance structure and other functioning rules in the Statutes. Its budget will be based on its members' contributions complemented by other sources of revenues, which may include EU and national grants. The seat of an EDIC is in a participating Member State and its legal personality must be recognised by all Member States.

An EDIC may implement a multi-country project by deploying joint infrastructure, delivering services and bringing together – as considered appropriate by the founding Member States – public entities, private entities, final users and industry. At the time of writing this roadmap, three EDICs were officially approved, being ALT-EDIC⁵⁴³ (EU infrastructure on Language Technologies to train large language models on EU's regional and official languages), CitiVERSE⁵⁴⁴ (connecting existing local Digital Twins of smart cities) and EUROPEUM-EDIC⁵⁴⁵ (developing the existing ecosystem of the European Blockchain Services Infrastructure).

32.3 Funding programs relevant for VHT

32.3.1 European programs

The European Commission has been funding many successful initiatives in the area of computer modelling and simulation in healthcare towards the realisation of Digital Twins in healthcare. To foster a thriving ecosystem around Digital Twins in healthcare, various public policy measures and funding instruments need to be implemented, such as research grants to support the development and testing of Digital Twin technologies, infrastructure development and operation investment, collaborative initiatives between public and private stakeholders, training and education programmes, pilot projects

⁵⁴² <https://digital-strategy.ec.europa.eu/en/policies/edic>

⁵⁴³ <https://alt-edic.eu/>

⁵⁴⁴ https://eur-lex.europa.eu/eli/dec_impl/2024/459/oj

⁵⁴⁵ <https://digital-strategy.ec.europa.eu/en/news/blockchain-creation-europeum-edic>

and testbeds, data sharing and interoperability, and incentives for adoption. and contribution Current funding for the development of Digital Twins in healthcare can be accessed under three main funding instruments in the field of healthcare, research and innovation established by the EC in 2021:

- **The Digital Europe Programme:** established via Regulation (EU) 2021/694⁵⁴⁶
- **Horizon Europe:** the framework programme for research and innovation, laying down its rules for participation and dissemination, established via Regulation (EU) 2021/695⁵⁴⁷
- **EU4 Health Programme:** established via Regulation (EU) 2021/522⁵⁴⁸

The **Digital Europe Programme** is designed to bring digital technology to businesses, citizens, and public administrations, while aiming to improve the EU's competitiveness in the global digital economy, contribute to bridging the digital divide across the EU and strengthen Europe's competences in digital technology through large-scale distribution. The Digital Europe Programme is running from January 2021 to December 2027, with an expected budget of €7.5 billion, under direct management of the European commission, and performs its work towards five specific goals:

- Achieving high-performance computing,
- Artificial Intelligence,
- Honing advanced digital skills,
- Optimising digital interoperability
- Advancing Cybersecurity.

A network of European Digital Innovation Hubs provides technological expertise for the participating entities (private and public). This initiative is co-financed by the Member States and is open to non-EU countries that are part of the EEA, candidate countries, European Neighbourhood Policy countries, and other countries under the agreement.

The 2023-2024 Digital Europe Work Programme described under Cloud, Data and Artificial Intelligence, addresses the **development of the Advanced Simulation Platform for advanced VHT models**⁵⁴⁹ (action 2.3.4). The objective of the action is to develop a distributed platform that provides access to a federated repository of VHT related resources, open source software toolkits, and computational services for developing, testing, and integrating VHT models. The platform will be used for personalised care, medical training, surgical intervention planning, and professional training and education purposes. It will provide controlled and secure access to an environment of simulation and visualisation tools, open access and proprietary data, and assets for advanced modelling. The platform will be interoperable with augmented and virtual reality (AR/VR) environments and will be based on access to computational services enabled by strategic digital capabilities with links to suitable testing and experimentation facilities. The indicative global budget is €20 million and the indicative duration of the action is 24-36 months⁵⁵⁰. *The ambition of the authors of this roadmap is to inform the development of this advanced simulation platform in terms of vision, tech stack, infrastructure needs and requirements, ELSI, standards, regulatory, as well as ecosystem incentives.* Other programs include the deployment of a pan-European federated infrastructure for ICU data.

Horizon Europe⁵⁵¹ is the key funding programme for Research and Innovation, acting for the period 2021 to 2027, with a total budget of €95.5 billion. Its overarching goal is to propel scientific, technological, economic, and societal advancement in the EU by investing in Research and Innovation. This programme intends to maximise added value by focusing on milestones that can only be achieved by the Member States acting in cooperation. The programme has several objectives, which include promoting scientific intelligence, supporting the creation and application of high-quality knowledge, skills, technologies, and solutions, providing training for researchers and supporting new talent. The program also aims to strengthen the impact of research and innovation in implementing Union policies, promote innovation in European industry and societal challenges, foster all facets of innovation, facilitate technology transfer and encourage excellence-based participation from all member states to

⁵⁴⁶ <https://eur-lex.europa.eu/EN/legal-content/summary/digital-europe-programme-2021-2027.html>

⁵⁴⁷ <https://eur-lex.europa.eu/eli/reg/2021/695/oj>

⁵⁴⁸ <https://eur-lex.europa.eu/EN/legal-content/summary/eu4health-programme-2021-2027.html>

⁵⁴⁹ <https://digital-strategy.ec.europa.eu/en/funding/platform-advanced-virtual-human-twin-vht-models>

⁵⁵⁰ <https://ec.europa.eu/newsroom/dae/redirection/document/94609>

⁵⁵¹ https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf

strengthen the attractiveness of the European Research Era. A range of calls is situated at the interface of the digital and the health focus. The first Strategic Plan covers the years 2021 – 2024.

The specific programme implementing Horizon Europe includes three pillars:

- Pillar I: Excellent Science, reinforcing and extending the excellence of the Union's science base (European Research Council expected budget €16 billion, Marie Skłodowska-Curie Actions - €66 billion and Research infrastructure -2.4 €billion).
- Pillar II: Global challenges & European Industrial Competitiveness, boosting key technologies and solutions underpinning the EU policies & Sustainable Development Goals. Expected funding for the health cluster is €8.2 billion.
- Pillar III: Innovative Europe, stimulating market-creating breakthroughs and ecosystems conducive to innovation, with expected budgets of 10.6 billion, including up to €527 million for European innovation ecosystems, and €3 billion for the European Institute of Innovation and Technology (EIT).

There are five Research and Innovation mission areas in Horizon Europe, which includes the EU **Mission on Cancer** that spans until 2030, which aims to support cancer patients living longer and better, achieve a thorough understanding of cancer, prevent what is preventable, optimise diagnosis and treatment, support the quality of life of all people exposed to cancer, and ensure equitable access to the above across healthcare. In addition, the Horizon Europe programme has established the European Innovation Council, which has a budget of ~€10b to support ground-breaking innovations throughout the lifecycle from early stage research, to proof of concept, technology transfer, and scale up of start-ups and SMEs in Europe.

Under this program, dedicated calls have supported the development of integrated, multi-scale computational models of patient patho-physiology for personalised disease management⁵⁵², as well as their integration with other technologies such as AR/VR to bring interactive visualisation tools and facilitate access to the available relevant data resources and models⁵⁵³.

The **EU4Health Programme**⁵⁵⁴ is the EU's response to the COVID-19 pandemic and has an expected budget of €5.3 billion during the 2021-27 period. It aims to support and complement national policies to improve and protect human health, strengthen health systems, and increase resource efficiency within the EU. In this light, this programme has 10 objectives:

- Disease prevention and boosting of health campaigns.
- Promotion of international health initiatives
- Efficient prevention and response to cross-border health threats
- Complementing national stockpiling of essential crisis-relevant products
- Creating a reserve of healthcare and support staff
- Improving availability and accessibility of medical products and devices
- **Strengthening digital tools and health data storage and transformation**
- Increasing access to healthcare
- Developing and implementing EU health legislation
- Collaboration between Member States' healthcare systems

Recent relevant calls under this program include one on advancing the adoption of AI in health.

32.3.2 Member state programs: some examples

The VHT initiative represents a transformative approach to healthcare innovation, leveraging cutting-edge Digital Twin technologies to enable personalised, predictive, and preventive care. The table given below outlines a comprehensive roadmap of funding opportunities at both the EU and Member State levels, specifically highlighting Belgium, Finland, France, the Netherlands, and Portugal. Additionally, it explores synergies with key EU projects such as QUANTUM, Digital Health Uptake (DHU), and xShare. These are some of the opportunities and collaborations that the VHT ecosystem can utilise to overcome sustainability challenges and increase its uptake. Together, these efforts may enhance the long-term sustainability and scalability of the VHT initiative across Europe.

⁵⁵² HORIZON-HLTH-2023-TOOL-05-03

⁵⁵³ <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-infra-2024-tech-01-04>

⁵⁵⁴ https://health.ec.europa.eu/funding/eu4health-programme-2021-2027-vision-healthier-european-union_en

Table 11: Funding opportunities for the VHT

Member State	Description	How can the VHT ecosystem leverage this opportunity
Belgium 	Health Data Agency Innovation Project Funding (2024) offered a targeted funding opportunity of €500,000 to support innovative projects that facilitate innovative use and reuse of health data , improving the quality, affordability, and accessibility of healthcare. Open to all entities, including associations, private organizations, public institutions, universities, and individuals in the health sector.	The Health Data Agency funding call is an opportunity for the VHT to get financial and strategic support in projects that advance health data reuse and integration - with the support of Belgian authorities.
Finland 	Health 360 Finland Programme is part of Finland's Healthcare Reimagined 2035 mission, offering a comprehensive platform to support innovative health solutions with its focus on promoting digital health, fostering international collaboration, and supporting exports. The programme targets Finnish companies of all sizes, especially SMEs, start-ups, and deep-tech firms as well as research organizations, universities, and foreign researchers, fostering collaboration for advanced healthcare innovation	The Health 360 program can provide support for the VHT, aligning its innovative health technology goals with Finland's strategic vision. Through networking, funding, and international collaboration opportunities, the program can strengthen the VHT's capacity to transform healthcare and expand into key global markets.
France 	MEDITWIN consortium for French organizations 2024-2029. This programme focuses on creating and deploying virtual twins for medical practice.	The MEDITWIN programme exemplifies how national investment in advanced digital health solutions can accelerate the adoption of VHT technologies. It fosters collaboration among leading research institutions, industrial leaders, and startups, MEDITWIN aligns well with the VHT's vision
The Netherlands 	HealthHolland encourages collaboration among businesses, universities, knowledge institutions, and government entities to share resources and expertise for faster innovations. HealthHolland coordinates national partnerships to align investments and activities for economic and societal benefits, guided by the Knowledge and Innovation Agenda Health & Care (2024-2027). It provides financial support for R&D through subsidies, enabling groundbreaking scientific discoveries in fields like eHealth, medical technology, and data analysis that align with VHT's focus.	HealthHolland provides a robust platform to advance VHT technologies while contributing to economic growth and societal well-being in the Netherlands and beyond.
Portugal 	The CaixaImpulse Health Innovation Program, led by the "la Caixa" Foundation, offers a unique opportunity to accelerate biomedical and health research projects. With a total funding of €3.4 million for the 2024 edition, the program supports initiatives in Spain and Portugal, aiming to transfer research results to society and the market.	The CaixaImpulse Health Innovation Programme gives a robust platform with financial support and collaboration opportunities to accelerate VHT technology development. With the facilitation of this program, the VHT initiative can expedite innovation, collaboration, and its mission to shape personalised medicine and healthcare in Europe.

These funding opportunities are examples of pathways for the VHT initiative to attain longer-term sustainability. By leveraging these, the relevant stakeholders are expected to find the initial resources to further develop the VHT infrastructure and to collaborate across Member States. Such efforts are expected to ensure early adoption and further scalability of VHT solutions while supporting the growth of a dynamic digital health ecosystem in Europe. *The ambition of the authors of this roadmap is to provide a clear vision,*

ambition and technology framework so that individual researchers, or local consortia, can develop their VHT projects in a way that is immediately interoperable with the EU VHT initiative. This ‘interoperability-by-design’ or ‘harmonisation-by-design’ will allow for important contributions to VHT realisation being made in a bottom-up fashion.

32.3.3 Public-private partnerships for VHT sustainability

European (Public-Private) Partnerships serve as a crucial mechanism for implementing Horizon Europe and advancing the European Research Area. The Horizon Europe Regulation introduced a new approach to **European Partnerships**⁵⁵⁵, emphasizing strategic coherence, impact-driven objectives, and systemic alignment. These partnerships are required to adhere to common criteria outlined in the Horizon Europe Regulation throughout their lifecycle, encompassing selection, implementation, monitoring, evaluation, and eventual phasing out.

European Partnerships are categorized into three forms: **co-funded, co-programmed, and institutionalised**. These partnerships aim to address critical EU policy priorities and facilitate the transformation of economies and societies, particularly through the green and digital transitions, while enhancing the resilience and competitiveness of European industries in response to global challenges and shifting market dynamics.

To maximize impact, European Partnerships are expected to adopt a systemic approach. This includes fostering synergies across research and innovation policies, improving sectoral integration, and ensuring inclusivity by being more accessible to newcomers. Additionally, partnerships must secure long-term commitments from both public and private partners and adopt a strategic outlook at the international level. Through these measures, European Partnerships aim to deliver transformative results that align with the EU's broader policy objectives.

Co-programmed partnerships, such as the **Partnership on AI, Data and Robotics (ADRA)**⁵⁵⁶ are formalised through Memoranda of Understanding, defining collaboration frameworks between the EU and public or private entities, with EU contributions funded through the Framework Programme. ADRA leverages 1.3 billion Euro of public investments through the Horizon Europe programme, complemented with 1.3 billion Euro of private investments in the period 2021-2030 to address the key challenges in European AI, Data and Robotics. ADRA's scope includes “All of AI”, “All of Data” and “All of Robotics” within its remit. Activities range from roadmapping and engagement, over early-stage research and innovation, to deployment in the private sector. The Partnership is open to industry, academia, public bodies and to organisations both small and large. In particular, SMEs, start-ups and entrepreneurs are encouraged to join the Partnership.

Co-funded partnerships such as the European Institute of Technology (EIT) – and its Knowledge and Innovation Center for Health, **EIT-Health**⁵⁵⁷ – is a strong, diverse and balanced European Partnership of best-in-class organisations in education, research, technology, business creation and corporate and social innovation. EIT Health's vision is ‘To enable people in Europe to live longer, healthier lives by building and growing businesses to create products and services that progress healthcare in Europe, while strengthening our economy and the sustainability of our healthcare systems.’ By 2030 EIT strives to be Europe's leading innovation platform, facilitating longer, healthier lives and more sustainable healthcare systems.

Another example of a co-funded partnership is the **Transforming Health and Care Systems partnership (THCS)**⁵⁵⁸. Health and care systems in Europe are facing core common challenges, which require harmonised and coordinated solutions. The European Partnership THCS represents a unique strategic opportunity to bring together stakeholders, create synergies, coordinate Research and Innovation actions, facilitate the digitisation of health and care services and support the transformation

⁵⁵⁵ Annex III of the REGULATION (EU) 2021/695 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013

⁵⁵⁶ https://projects.research-and-innovation.ec.europa.eu/sites/default/files/bmr-2022/ec_rtd_bmr-2022-ai-data-and-robotics-fiche.pdf and <https://adr-association.eu/>

⁵⁵⁷ https://projects.research-and-innovation.ec.europa.eu/sites/default/files/bmr-2022/ec_rtd_bmr-2022-eit-health-fiche.pdf and <https://eithealth.eu/>

⁵⁵⁸ <https://cordis.europa.eu/project/id/101095654/reporting> and <https://www.thcspartnership.eu/>

of health and care systems with innovative solutions driven by knowledge and evidence. The general objective of THCS is to contribute to the transition towards more sustainable, efficient, resilient, inclusive, innovative and high-quality people-centred health and care systems equally accessible to all people.

An example of an institutionalised partnership is the **Innovative Health Initiative** Joint Undertaking (IHI JU⁵⁵⁹). IHI JU is a public-private partnership between the European Union, represented by the European Commission, and several health industries from the biopharmaceutical, biotechnology and medical technology sectors, represented by their trade organisations (MedTechEurope, COCIR, EFPIA, EuropaBio). By leveraging the skills of a range of scientific fields - from pharmaceuticals and biotechnology to medical technology and big data - IHI funded projects that will address public health needs, improve patients' lives, and boost the competitiveness of Europe's health industries. In the past years, several VHT-related actions have been supported on comprehensive stroke management with predictive computational models, integrated patient health data, and improved visualisation.

32.4 Towards an established VHT public infrastructure

To establish a durable and scalable foundation for the VHT, a VHT-EDIC should be established. Through an EDIC, the VHT can integrate with existing European initiatives while ensuring interoperability, financial security, and stakeholder inclusivity. Upon reviewing the applications of the founding Member States, the European Commission will issue an implementing decision to establish the European Digital Infrastructure Consortium. The decision is made in consultation with the Digital Decade Policy Programme Committee and is subject to meeting all the requirements outlined in the Digital Decade Policy Programme 2030⁵⁶⁰.

The EDIC framework offers a **legal and governance structure that** ensures equitable and transparent decision-making within participants. Founding Member States collaborate to define governance rules, including the allocation of voting rights and the distribution of resources. This participatory approach minimises asymmetries between smaller and larger Member States, fostering collaboration. For example, the European Blockchain Services Infrastructure (EBSI)⁵⁶¹, which uses the EDIC model, has demonstrated how shared governance can streamline decision-making while maintaining equitable participation across nations. For the VHT, the framework can provide:

- **Inclusivity of Member States:** by enabling smaller and larger Member States to collaborate under the same governance framework, the EDIC model helps contributors to have a voice. This prevents dominance by larger nations and fosters balanced representation, critical for gaining widespread support and participation of the VHT.
- **Broad stakeholder engagement:** the framework extends beyond governments to include industry players, research institutions, and patient advocacy groups, creating a diverse ecosystem that is essential for tackling the complexities of healthcare innovation.

For the VHT, this inclusivity would ensure that the infrastructure is co-designed with input from a wide range of perspectives, resulting in a system that reflects the needs of all stakeholders, including underrepresented groups or less-developed healthcare systems in different Member States.

A cornerstone of the EDIC framework is its ability to **combine diverse funding streams**. Contributions from Member States are augmented with EU grants, such as those available under the Horizon Europe program, and private investments. This blended funding approach significantly reduces financial risk and creates opportunities for sustained resource allocation. Studies on the European High-Performance Computing (EuroHPC) Joint Undertaking show that diversified funding models have enabled the development of critical infrastructure, supporting innovation and cross-border collaboration in computational tasks⁵⁶². These optimisations are interesting for VHT sustainability as:

⁵⁵⁹ https://projects.research-and-innovation.ec.europa.eu/sites/default/files/bmr-2022/ec_rtd_bmr-2022-innovative-health-initiative-fiche.pdf and <https://www.ihj.europa.eu/>

⁵⁶⁰ European Commission, 'European Digital Infrastructure Consortium (EDIC) | Shaping Europe's Digital Future', accessed 9 October 2024, <https://digital-strategy.ec.europa.eu/en/policies/edic>.

⁵⁶¹ EBSI, 'About Us', accessed 9 October 2024, <https://ec.europa.eu/digital-building-blocks/sites/display/EBSI/About+us>.

⁵⁶² Thomas Skordas, 'Toward a European Exascale Ecosystem: The EuroHPC Joint Undertaking', Communications of the ACM 62, no. 4 (20 March 2019): 70–70, <https://doi.org/10.1145/3312567>.

- Diversified funding **reduces long-term adoption risks**: the EDIC framework combines Member State contributions, EU grants, and private investments, ensuring financial stability and reducing reliance on any single funding source.
- **Sustained resource allocation**: this blended model provides consistent funding to maintain and scale VHT infrastructure over time, supporting its long-term viability.
- **Proven success**: similar approaches, like the European HPC Joint Undertaking, have successfully fostered innovation and cross-border collaboration, demonstrating the effectiveness of diversified funding for large-scale projects.

The EDIC framework plays a role in **ensuring the VHT aligns seamlessly with existing EU initiatives**. This alignment may foster interoperability, which is the ability of systems and platforms to work together by integrating and exchanging data efficiently. **Interoperability** is particularly important for the VHT as it enables cross-border collaboration, allowing diverse data streams from different Member States to be integrated into a unified and functional system.

The European Health Data Space (EHDS) serves as a similar model, providing standardised data formats and secure mechanisms for sharing sensitive information across countries. Incorporating the principles and standards from initiatives like EHDS, the VHT can ensure that its infrastructure supports seamless data exchange, enhances collaboration between stakeholders, and adheres to robust privacy and security frameworks. This compatibility makes the EDIC framework as an excellent option for the VHT's long-term scalability and effectiveness across Europe.

33 Users, uptake and sustainability for VHT: conclusions and recommendations

33.1 Conclusions

Part 5 of the roadmap delves into the crucial aspects of engaging users, promoting the adoption of the Virtual Human Twin VHT, and ensuring its long-term sustainability. This part emphasizes the need for a collaborative and inclusive approach that involves all stakeholders, from researchers and clinicians to patients and policymakers.

A thriving and engaged VHT ecosystem is essential for the success of the initiative. Cultivating a strong community that embraces the VHT vision and actively contributes to its development requires understanding the **incentives** for both VHT creators and consumers. Researchers and model developers are motivated by the prospect of accessing vast datasets, validating their models, and collaborating with a global community. Clinicians, on the other hand, seek demonstrable clinical benefits, improved patient outcomes, and enhanced decision-making tools. Addressing the specific needs and expectations of each stakeholder group is paramount for fostering their engagement and maximizing the impact of the VHT.

Developing **sustainable business models** for the VHT is crucial for ensuring its long-term viability and continued development. Identifying potential revenue streams, exploring different funding mechanisms, and establishing a robust governance framework will enable the VHT to operate sustainably and evolve to meet the changing needs of its users. The roadmap explores various business model strategies, including public funding, private investment, subscription models, and data licensing agreements, highlighting the importance of a phased approach that adapts to the evolving landscape of the VHT ecosystem.

Establishing a VHT Marketplace is envisioned as a key facilitator for connecting VHT creators and consumers, fostering collaboration, and promoting the uptake of VHT technologies. The marketplace will serve as a platform for sharing and accessing data, models, and computational resources, streamlining the development and deployment of VHT applications. By providing a centralized hub for VHT-related activities, the marketplace will accelerate innovation, stimulate economic growth, and drive the adoption of the VHT across various healthcare domains.

A robust, publicly funded infrastructure is vital in realizing the full potential of the VHT. This infrastructure should leverage existing European initiatives and infrastructures, while addressing the specific needs and requirements of the VHT. A **dedicated VHT infrastructure** and diverse funding landscape are required to guarantee sustainability. The successful incorporation of this infrastructure in a European Digital Innovation Consortium requires a collaborative effort involving the European Commission, EU member states, research institutions, healthcare organisations and industry stakeholders.

33.2 Recommendations

Ensuring the long-term **sustainability** of the VHT requires a multifaceted approach that encompasses financial, technical, social, and ethical considerations. The roadmap identifies key factors for achieving sustainability, including securing stable and diverse funding sources, developing robust governance mechanisms, fostering community engagement, promoting open-source principles, and addressing the ethical and social implications of the VHT. A collaborative effort involving all stakeholders is essential for building a sustainable VHT ecosystem that delivers on its promise to transform healthcare. As such, following recommendations can be made.

33.2.1 Users & inclusiveness

1. **Creating an inclusive VHT ecosystem:** For the Virtual Human Twin to successfully integrate into healthcare systems, it is crucial to foster an active and outgoing ecosystem that embraces

diverse perspectives. This involves proactively gathering feedback from all stakeholders, including citizens, patients, industry professionals, healthcare providers, and scientists into the VHT development, design and materials. Co-creation with societal partners, raising awareness of the VHT's potential, and empowering users through accessible support mechanisms will build trust and promote wider adoption. Embracing innovative educational approaches will be key to fostering a broader understanding of the VHT, while leveraging advanced technologies, such as extended reality and AI, will facilitate user interactions. A robust support system encompassing access for all, dedicated researcher support, user-friendly tools like wizards, mapping tools, as well as dedicated training and workshops will further strengthen the VHT ecosystem.

2. **Promoting clinical and patient uptake of VHT:** To ensure the successful adoption and integration of VHT technologies into clinical practice, it is essential to prioritize stakeholder engagement and address the needs of clinical end-users throughout the development cycle. Co-creation with clinicians and patients, combined with well-defined incentives, can build trust and foster collaboration. These incentives should focus on delivering clear clinical benefits such as addressing unmet needs, improving patient outcomes, reducing costs, and enhancing care quality. Practical challenges, including data privacy, legal and ethical concerns, and cost implications, must be systematically addressed with robust mechanisms and transparent communication. Additionally, inclusion in clinical guidelines and offering tools and services aligned with clinical workflows, such as seamless integration with hospital systems and (immersive) patient engagement tools can further demonstrate the VHT's relevance. Outreach efforts, including educational programs and representation at clinical events, are vital to increase awareness and engagement. Finally, empowering clinicians through leadership roles and incorporating patient-centred design principles will strengthen trust and ownership, fostering widespread acceptance and accelerating the VHT's integration into healthcare systems.
3. **Ensuring equitable access to the VHT:** The Virtual Human Twin should be accessible to individuals of all ages, genders, ethnicities, socioeconomic statuses, and disabilities. This will require fostering digital literacy and promoting equitable access to high-quality healthcare. To ensure the VHT is representative of diverse populations, it is crucial to incorporate data from rare diseases, special populations, ageing populations, and individuals residing in developing countries. Access to VHT technology will empower individuals to manage their health through personal health forecasting, improving health outcomes for all. Immersive technologies coupled to VHT applications can play an important role to ensure full appreciation of VHT results and lower barriers to adoption,
4. **Communicating the VHT - reaching diverse audiences:** The Virtual Human Twin initiative requires a multifaceted communication strategy to reach diverse audiences across the European Union. Information should be disseminated in all EU languages to ensure inclusivity and accessibility. Engaging science communication experts can help translate complex scientific concepts into easily understandable language for the public. A combination of communication channels, including blogs, social media, science fairs, and interactive events like "science nights" will foster broader public awareness and understanding of the VHT. Developing user-friendly visualization tools will assist VHT developers and users in making the obtained results more accessible and engaging for a wider range of stakeholders. Citizen science initiatives can also be incorporated to promote public participation and engagement in VHT research.
5. **Training the future VHT workforce:** To fully realize the potential of the Virtual Human Twin in transforming healthcare, comprehensive education and training programs are essential for all stakeholders. This includes incorporating training on digital technologies, including AI and computational modelling, into all healthcare-related degrees. Existing healthcare professionals and providers require retraining to acquire foundational knowledge of VHTs and their practical applications, enabling them to confidently work with this emerging technology. This widespread upskilling will foster a workforce capable of effectively contributing to and benefiting from VHT advancements, facilitating its seamless integration into clinical practice. As the VHT evolves, new professional roles could emerge focused on liaising between organisations and the VHT, such as VHT innovation officers.

33.2.2 Sustainability

6. **Incentivising the uptake of a credible VHT platform:** To ensure the sustainability of the VHT platform, it is crucial to incentivise its adoption in new developments while promoting the reuse of existing efforts, including data and models. This can be achieved through establishing clear governance mechanisms and implementing incentives for sharing resources such as harmonisation and standardisation tools, and promoting international cooperation to broaden the VHT knowledge base. Developing measures and metrics to track and reward resource sharing, evaluating model quality, and emphasizing fairness and accessibility will further support the platform's long-term maintenance, longevity, and affordability.
7. **Stimulate VHT commercialisation:** To accelerate the development of commercial activities around the VHT, establishing a robust marketplace is essential. This should provide a range of options to facilitate the exchange of VHT-related resources while ensuring the protection of IP rights. Developing diverse and sustainable business models will ensure long-term scalability and growth. Thoroughly understanding the business case of future adopters, including their willingness to pay for VHT services, is essential. Drawing insights from existing commercial infrastructures can provide valuable guidance in shaping the VHT marketplace and business models, and foster the VHT's successful integration into the healthcare ecosystem.
8. **Investment strategy for the VHT:** To ensure the successful development and implementation of the VHT, continuous, robust and diverse sources of support; both public and private, are essential. This support will facilitate a wide range of activities from existing platform maintenance, over the integration of new models and modalities, to ensuring legal compliance. A long-term investment strategy by the EU, including new EU calls for proposals, will attract multi-funder contributions and encourage co-investment from Member States and private partners in developing the VHT infrastructure and applications. Moreover, securing long-term funded staff and promoting resource reuse will further contribute to the sustainability and advancement of the VHT ecosystem.
9. **Integrating the VHT into the European infrastructure landscape:** European policymakers should champion policy innovation to accelerate VHT development and adoption. The VHT's ambitious scope requires its integration into the European Infrastructure landscape. A decentralized approach is recommended, establishing national nodes to promote collaboration and resource sharing across Europe (*cfr.* EOSC). Establishing the VHT as a European Digital Infrastructure Consortium or forming a dedicated European Partnership for the VHT will provide a sustainable framework for its long-term development, implementation and ecosystem engagement. This necessitates the active involvement of EU Member States as well as public and private partners to pool resources, expertise, and infrastructure, fostering collaboration across national borders.
10. **The importance of environmental sustainability for the VHT:** The development and deployment of the VHT needs to include careful consideration of its environmental impact. Minimising the environmental footprint of the VHT infrastructure requires examining the energy consumption associated with data storage, processing, and model execution. This necessitates adopting sustainable practices in data centre operations and exploring energy-efficient algorithms and hardware architectures. Embracing the principles of Planetary Health, the VHT should strive to contribute to a healthier planet, recognizing the interconnectedness between human health and the environment. This underscores the importance of applying a "sustainability triangle" approach to VHT development, encompassing environmental, social, and economic considerations to ensure its long-term viability.

PART 6:

RECOMMENDATIONS FOR A SUCCESSFUL VHT

34 Recommendations for the Roll-Out of the Virtual Human Twin

34.1 A tentative timeline of activities

This chapter outlines recommendations for the successful roll-out of the Virtual Human Twin (VHT) over the next decade, emphasizing key areas crucial for its realisation. The recommendations acknowledge the need for synergies with and dependencies on other initiatives, encompassing research, legislation, regulatory frameworks, and ethical dimensions, to ensure the VHT's effective and responsible integration into the healthcare landscape.

The recommendations discussed in the next sections are frequently inter-dependent and will only be realized through a gradual roll-out over time. In the figure below, a proposal is made to pace the work for **realising the VHT across the coming 10 years**, divided into the following 3 phases:

- phase 1 (2025-2028, short term): building the foundations and demonstrating potential
- phase 2 (2029-2031, middle-long term): integration and continued investment
- phase 3 (2032-2034, long term): deployment (in the hospital).

Several of the activities listed in the early phases, such as ecosystem engagement, communication, policy watch and use case development, will require **sustained efforts throughout** the later phases too. The work indicated in a particular phase is contingent on the successful realization of the work in earlier phases (milestones), such as the presence of a functioning VHT simulation platform at the start of phase 2. At the same time, there are dependencies to a range of developments that are not strictly related to the VHT but are relevant to the different elements, including technologies (*e.g.* AI, sensors), infrastructure (EHDS and other infrastructures) and ELSI (acts, regulations). These dependencies might expedite or delay the roll-out of certain activities mentioned in the recommendations.

Many of the recommendations and dependencies outlined in the following sections are expected to be effectively facilitated through the EOSC EU node, launched in October 2024:

- **Access to High-Performance Computing (HPC) Resources:** The EOSC EU node provides access to state-of-the-art HPC resources, including high-powered servers and specialized software.
- **Interoperability and Data Sharing:** By promoting FAIR principles (Findable, Accessible, Interoperable, Reusable), the EOSC EU node enhances interoperability and facilitates seamless data sharing.
- **Advanced AI and Machine Learning Tools:** The platform offers cutting-edge AI and machine learning tools that can be used to train VHT models, analyse large datasets, and generate predictions related to human health and disease.
- **Secure and Reliable Infrastructure:** Built on a robust and scalable foundation, the EOSC EU node ensures secure and reliable infrastructure for storing and processing sensitive VHT data, safeguarding patient privacy and maintaining data integrity.

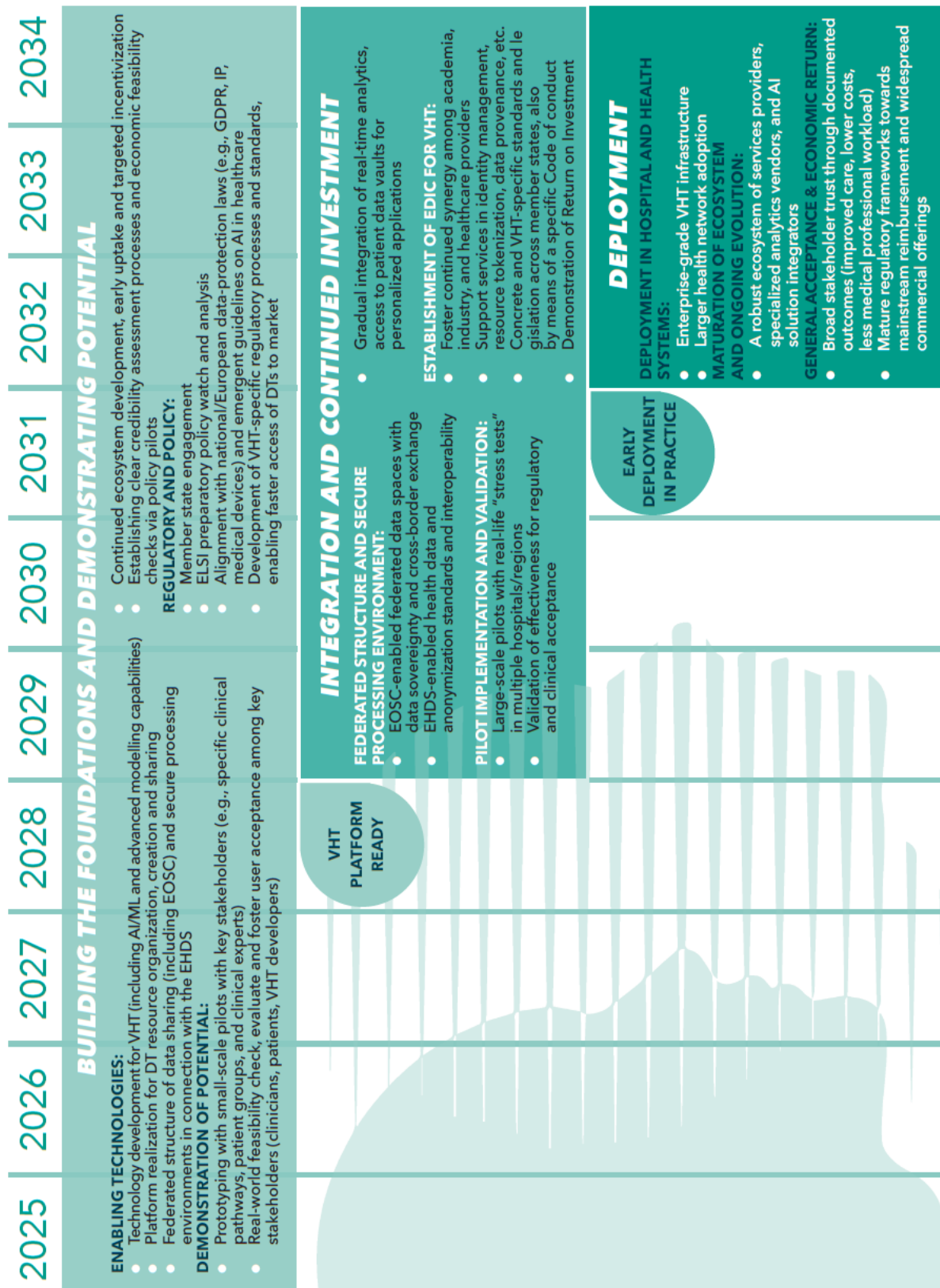


Figure 46: Schematic representation of key activities for VHT over time.

This chapter outlines recommendations for the successful roll-out of the Virtual Human Twin (VHT) over the next decade, emphasizing key areas crucial for its realization. The recommendations acknowledge the need for synergies with and dependencies on other initiatives, encompassing research,

legislation, regulatory frameworks, and ethical dimensions, to ensure the VHT's effective and responsible development and integration into the healthcare landscape.

34.2 Need assessment of creators and consumers

1. **Research and innovation:** to ensure that the Virtual Human Twin reaches its full potential, research and innovation (R&I) are paramount. This R&I should span a continuum from basic research, addressing fundamental knowledge gaps and the development of nascent technologies, to translational research, bridging the gap between laboratory findings and clinical applications. The scope of R&I should cover the creation of generic and population-specific Digital Twins as well as highly personalized Digital Twins, catering to diverse applications. Embracing blue-sky research will be crucial for breakthroughs and innovation, alongside the exploration of diverse use cases to demonstrate the VHT's broad applicability.
2. **Prioritising use cases to enhance VHT impact:** the identification, development, and delivery of high-impact validated use cases is important to demonstrate the practical value of the VHT across diverse clinical and scientific applications. Prioritization should be given to use cases that address key areas such as diagnostics, medical education and training, clinical decision support, therapy development, and intervention planning. The development of a demonstrator (such as the Virtual Brain Twin) showcasing the capabilities of the VHT in a specific area will be valuable for engaging stakeholders and showcasing its potential. Selecting a diverse range of use cases will illustrate the VHT's versatility and broad applicability, particularly highlighting the transition from screening diseases to more complex stratified simulations. This approach ensures that the VHT is developed with a focus on real-world applications, maximising its impact and adoption.
3. **Enhancing clinical usefulness - a cornerstone for VHT adoption:** to ensure the VHT is embraced in clinical practice, demonstrating its clinical usefulness is paramount. This requires a rigorous assessment of its clinical effectiveness and usability. Following a co-creation approach, the conceptualisation and development of the VHT must be driven by a clear understanding of clinical needs, ensuring the VHT addresses practical challenges and provides tangible benefits. Fostering dialogue and collaboration with clinicians throughout the development process will ensure the VHT is aligned with clinical workflows and priorities.
4. **Applications across the disease continuum - realizing the full potential of the VHT:** to maximise the impact of the Virtual Human Twin, it is essential to advance the understanding of how its solutions, products, and services can be applied across the entire disease continuum. This includes exploring its potential in prevention, treatment, and follow-up, spanning a wide range of applications from biomedical and clinical studies to therapy development and diagnostics. The VHT should not be confined to traditional healthcare settings but also encompass innovative care models such as remote care and self-care, empowering individuals to actively manage their own health. This comprehensive approach will unlock the full potential of the VHT, transforming healthcare from a reactive system to a proactive, personalised, and patient-centric model.

34.3 VHT technologies: basic building blocks

5. **Advancing *in silico* technologies for a robust, versatile and impactful VHT:** to fully realise the transformative potential of the VHT, substantial investment in the development of *in silico* technologies is crucial. This includes focusing on longitudinal and longer-term human health and disease modelling, capturing the long-term dynamic nature of human physiology. The development of multiscale models is essential for integrating knowledge across different levels of biological organization, from cells to organs to the whole body. Incorporating new data modalities beyond traditional clinical data will enrich the VHT, enhancing its predictive power. Omics integration will provide a comprehensive view of individual biological processes, enabling personalized insights into health and disease. AI will play a central role in the VHT, not only in data analysis and model generation but also in driving the interaction between AI and mechanistic modelling.
6. **Data-generating hardware addressing a critical need for VHT:** the development of advanced data-generating hardware technologies is essential for the progress of VHT applications. Fit-for-modelling-purposes, disease-specific data is currently lacking from traditional sources like hospitals. This data gap can be addressed by investing in wearable and implantable sensor

technologies capable of providing real-time patient data. Micro-physiological system technologies present another avenue for generating high-quality human-relevant data by combining microfluidics with embedded sensors. These technologies provide a platform for investigating adverse drug effects and modelling specific organ functions, offering valuable insights for the development of VHTs.

7. **Evidence generation for VHT development:** the development of robust and trustworthy VHT technologies hinges on the generation of comprehensive clinical, experimental, and digital evidence to support model building and validation. This process requires well-designed studies and clinical trials that first assess the technical validity of VHT solutions, followed by their clinical domains such as safety, effectiveness, and usability. A deep understanding of the unmet needs of VHT early adopters and how these technologies will be implemented to address those needs is essential for generating relevant and impactful evidence. VHT technologies can be used to develop synthetic and open-access validation datasets, ensuring transparency, reproducibility, and accessibility throughout this process. This evidence generation process will be instrumental in building confidence among stakeholders, paving the way for wider adoption of VHT in healthcare
8. **Credibility is key to building trust in VHT technologies:** to foster trust and confidence in VHT technologies, establishing their credibility is essential. This involves a multi-faceted approach encompassing rigorous and transparent safety assessments and peer review/validation of methods to ensure scientific rigor and transparency. Clear procedures for verification and validation, for a well-defined question of interest and context of use, must be developed, drawing upon established ontologies to provide a standardized framework for credibility assessment. Uncertainty quantification plays a vital role in understanding the limitations of VHT predictions and enabling risk-informed decision-making. Addressing potential refutations and providing clear evidence to support the credibility of VHT findings will be crucial for gaining acceptance from stakeholders.

34.4 VHT infrastructure

9. **Designing the VHT resource repository and simulation platform:** the design, construction, and continuous enhancement of the VHT resource repository and simulation platform are crucial aspects. It is imperative that this platform adheres to all applicable laws and regulations in Europe, particularly concerning data privacy and security. The platform should leverage distributed and federated computing, incorporating regional or country-specific nodes to ensure data sovereignty and efficient resource utilisation. A federated authentication and authorization mechanism is essential to manage access securely and facilitate collaborative data sharing across the VHT ecosystem. In addition to security, the platform should prioritise interoperability, modularity, flexibility, scalability and extensibility to accommodate future growth and advancements in VHT technologies. Embracing the principles of Open Science, Open Source, and Open Data will foster collaboration, accelerate innovation, and enhance transparency.
10. **Investing in a robust and adaptable IT infrastructure:** to ensure the longevity and continuous growth of the VHT, sustained support for the development, testing, and implementation of advanced and interoperable IT platform architectures is needed. This includes investments in computational infrastructure, cybersecurity, HPC, cloud services, and edge infrastructure to store, manage and process the vast amounts of data generated by the VHT. Continuous inclusion and updates of domain-specific services are crucial to enable advanced workflows and applications, provide tools for seamless integration, harness new technologies, and visualise results.
11. **The Future of the VHT - Data Availability and Access:** for the VHT to reach its full potential, access to high-quality, annotated, and interoperable digital health data is paramount. The VHT initiative should work in conjunction with existing and developing infrastructures and platforms, like the European Health Data Space. Standardised formats and semantics will be essential for interoperability and effective data analysis. Patient privacy, personal data, health, and safety must be safeguarded, which can be achieved through anonymisation, pseudonymisation, and synthetic data generation. Data curation, certification, and quality control will ensure the reliability of the data, with data stewards and curators playing key roles.
12. **Co-evolution of the VHT - a collaborative ecosystem for success:** the VHT should not exist in isolation but actively co-evolve with existing digital services and research infrastructures. Embracing a global vision for the European life sciences research infrastructures, the VHT

initiative should establish strong links to existing platforms, including ESFRI and EOSC, to avoid duplication and foster synergy. The VHT should leverage existing efforts and available platforms, mapping existing modelling and computational infrastructures, both commercial and open source, to ensure efficient resource utilisation and maximise impact. Long-term support is crucial, requiring engagement from national networks and continuous collaboration with other initiatives to create a thriving ecosystem for the advancement of the VHT.

13. **Fostering collaboration - Co-designing the VHT for optimal usability:** to ensure the successful adoption of the VHT technology, a co-design approach including end-users and other stakeholders is essential. Interactive design sessions can foster inclusion, allowing stakeholders to provide feedback and contribute to the development of a user-friendly platform. This user-centric design places ease of use at the forefront through documentation, demos, and training materials tailored to specific user needs. A strong emphasis on user experience and usability testing will ensure the platform's functionality aligns with the needs and expectations of diverse stakeholders. The development of an automated one-stop-shop catalogue can streamline access to resources and best practices, reducing barriers between research and practice. This participatory approach emphasises communication and transparency by using clear language to communicate design processes, composition, and best practices to user groups.

34.5 ELSI, standards & regulatory

14. **Enabling a robust regulatory landscape for VHT:** to effectively enable the efficacy, safety, trustworthiness, performance, and risk management of the VHT from its early stages of development, it is crucial to enhance the clarity of the regulatory landscape. This can be achieved through a credibility-by-design evolutionary framework that prioritises the establishment of clear standards, approaches, tools, and techniques. Key elements include the implementation of a regulatory assist framework, training and verification labs, and resources to assist in the creation of technical dossiers for certification. Robust validation and verification infrastructure is critical, as is the ability to assess data and model credibility. Ensuring data accuracy for validation and providing clear information about data accuracy will be crucial for building trust and ensuring the reliability of VHT outcomes. The establishment of dedicated sandboxes are also recommended to encourage innovation and accelerate the adoption of VHT technology.
15. **Prioritizing standards for a unified and interoperable VHT:** the VHT will heavily rely on standardization to guarantee interoperability. This includes well-defined metadata standards, consistent terminology, and efficient input procedures for quality assurance. Embracing ISO standards, unifying existing standards where applicable, and prioritising standards-based interoperability will ensure the VHT's credibility and widespread adoption. Establishing best practices and consensus procedures, particularly within mature communities that lack fixed standards, will be crucial. The development of clear guidelines for reporting and development, including standard operating procedures, will promote best practice sharing. This approach will ensure data can be seamlessly integrated and modelling results effectively validated.
16. **Navigating the legal landscape: ensuring compliance and adoption within the VHT:** to ensure the successful implementation of the VHT within the European Union, proactive monitoring and harmonisation of relevant EU and national legislation is necessary. The VHT platform should be designed to facilitate legal compliance, incorporating features and tools that help users identify and adhere to applicable regulations. This involves developing a clear accountability framework that outlines liability, provides reassurance, and establishes responsibility through guidelines and a code of conduct for VHT users. It is crucial to identify and address challenges within existing European legislation (*e.g.* AI Act, MDR, GDPR) that might hinder the certification and adoption of VHT solutions in industry and clinical settings. A unified EU-wide approach, or at the very least, careful monitoring of national laws and recognition of differentiation across Member States will be necessary to promote legal certainty and ensure seamless collaboration.
17. **IP and the VHT, a foundation for collaboration:** establishing common ground regarding IP Rights management and the protection of trade secrets is crucial for fostering trust and collaboration among VHT stakeholders. Additionally, the implications of software license's choices should be carefully evaluated to balance the need for open access with the protection of IP, an element that is important to facilitate the uptake of the platform and its application by start-

ups and larger industrial players. Licensing frameworks operating under Fair, Reasonable, and Non-Discriminatory terms, would ensure broad accessibility while compensating innovators. A comparable approach could be applied to the VHT ecosystem where specific foundational AI tools or methods are essential for interoperable and scalable systems.

18. **The role of health technology assessment in VHT - ensuring value and efficiency:** innovation adoption instruments, like pre-commercial procurement, innovation procurement, innovation partnerships and value-based procurement, enable evidence generation, the analysis of the impact of VHT on clinical cost structures and potential reimbursement pathways. A comprehensive innovation adoption process will facilitate translation of VHT research into practice. This includes pro-actively engaging with payers to demonstrate the synergies, value and cost-effectiveness of VHT applications across the value chain. Harmonizing approaches across VHT adopters and Member States will be key to unlocking the competitive advantage and leadership potential of the European VHT initiative and pave the way for its wider adoption, improving healthcare outcomes and promoting sustainable healthcare systems.
19. **Building trust in VHT - ethical considerations:** the future of the VHT requires sustained efforts to build trust among users, recognizing the ethical complexities of this emerging technology. While VHT offers strong potential to increase fairness and equality in healthcare, guaranteeing digital human rights and equitable access is paramount and will constantly need to be attentively pursued. Data ownership and models, including the role of altruism, will require careful consideration to ensure responsible development and utilization of the VHT. Continued research and open discourse on the ethical challenges and principles related to VHT, such as identity, human enhancement and technology dependence, are essential. A multi-disciplinary ethic-by-design process will ensure ethical considerations are fully embedded within the design and operation of the VHT.
20. **Embedding responsible research and innovation in the VHT:** to ensure the responsible development of the Virtual Human Twin, the genuine collaboration with Social Sciences and Humanities is crucial. This will increase stakeholder inclusion and reflexivity within the ecosystem. In addition, it will provide valuable insights into the societal impact of VHT, ethical considerations, and potential challenges related to user adoption and trust. Effective communication and dissemination strategies are essential, tailored to the specific needs and interests of various stakeholder categories. Bringing together experts from diverse fields will facilitate a holistic understanding of the complex technical, social, and ethical aspects of VHT development across the ecosystem.

34.6 Users & inclusiveness

21. **Creating an Inclusive VHT Ecosystem:** For the Virtual Human Twin to successfully integrate into healthcare systems, it is crucial to foster an active and outgoing ecosystem that embraces diverse perspectives. This involves proactively gathering feedback from all stakeholders, including citizens, patients, industry professionals, healthcare providers, and scientists into the VHT development, design and materials. Co-creation with societal partners, raising awareness of VHT potential, and empowering users through accessible support mechanisms will build trust and promote wider adoption. Embracing innovative educational approaches will be key to fostering a broader understanding of the VHT, while leveraging advanced technologies, such as extended reality and AI, will facilitate user interactions. A robust support system encompassing access for all, dedicated researcher support, user-friendly tools like wizards, mapping tools, as well as dedicated training and workshops will further strengthen the VHT ecosystem.
22. **Promoting Clinical and Patient Uptake of VHT:** To ensure the successful adoption and integration of VHT technologies into clinical practice, it is essential to prioritise stakeholder engagement and address the needs of clinical end-users throughout the development cycle. Co-creation with clinicians and patients, combined with well-defined incentives, can build trust and foster collaboration. These incentives should focus on delivering clear clinical benefits such as addressing unmet needs, improving patient outcomes, reducing costs, and enhancing care quality. Practical challenges, including data privacy, legal, social and ethical concerns, and cost implications, must be systematically addressed with robust mechanisms and transparent communication. Additionally, inclusion in clinical guidelines and offering tools and services

aligned with clinical workflows, such as seamless integration with hospital systems and (immersive) patient engagement tools can further demonstrate the VHT's relevance. Outreach efforts, including educational programs and representation at clinical events, are vital to increase awareness and engagement. Finally, empowering clinicians through leadership roles and incorporating patient-centred design principles will strengthen trust and ownership, fostering widespread acceptance and accelerating the VHT's integration into healthcare systems.

23. **Ensuring equitable access to the VHT:** the Virtual Human Twin should be accessible to individuals of all ages, genders, ethnicities, socioeconomic statuses, and disabilities. This will require fostering digital literacy and promoting equitable access to high-quality healthcare. To ensure the VHT is representative of diverse populations, it is crucial to incorporate data from rare diseases, special populations, ageing populations, and individuals residing in developing countries. Access to VHT technology will empower individuals to manage their health through personal health forecasting, improving health outcomes for all. Immersive technologies coupled to VHT applications can play an important role to ensure full appreciation of VHT results and lower barriers to adoption,
24. **Communicating the VHT - reaching diverse audiences:** the Virtual Human Twin initiative requires a multifaceted communication strategy to reach diverse audiences across the European Union. Information should be disseminated in all EU languages to ensure inclusivity and accessibility. Engaging science communication experts can help translate complex scientific concepts into easily understandable language for the public. A combination of communication channels, including blogs, social media, science fairs, and interactive events like "science nights" will foster broader public awareness and understanding of the VHT. Developing user-friendly visualization tools will assist VHT developers and users in making the obtained results more accessible and engaging for a wider range of stakeholders. Citizen science initiatives can also be incorporated to promote public participation and engagement in VHT research.
25. **Training the future VHT workforce:** to fully realize the potential of the Virtual Human Twin in transforming healthcare, comprehensive education and training programs are essential for all stakeholders. This includes incorporating training on digital technologies, including AI and computational modelling, into all healthcare-related degrees. Existing healthcare professionals and providers require retraining to acquire foundational knowledge of VHTs and their practical applications, enabling them to confidently work with this emerging technology. This widespread upskilling will foster a workforce capable of effectively contributing to and benefiting from VHT advancements, facilitating its seamless integration into clinical practice. As the VHT evolves, new professional roles could emerge focused on liaising between organisations and the VHT, such as VHT innovation officers.

34.7 Sustainability

26. **Incentivising the uptake of a credible VHT platform:** to ensure the sustainability of the VHT platform, it is crucial to incentivize its adoption in new developments while promoting the reuse of existing efforts, including data and models. This can be achieved through establishing clear governance mechanisms and implementing incentives for sharing resources such as harmonization and standardization tools, and promoting international cooperation to broaden the VHT knowledge base. Developing measures and metrics to track and reward resource sharing, evaluating model quality, and emphasizing fairness and accessibility will further support the platform's long-term maintenance, longevity, and affordability.
27. **Stimulate VHT commercialisation:** to accelerate the development of commercial activities around the VHT, establishing a robust marketplace is essential. This should provide a range of options to facilitate the exchange of VHT-related resources while ensuring the protection of IP rights. Developing diverse and sustainable business models will ensure long-term scalability and growth. Thoroughly understanding the business case of future adopters, including their willingness to pay for VHT services, is essential. Drawing insights from existing commercial infrastructures can provide valuable guidance in shaping the VHT marketplace and business models, and foster the VHT's successful integration into the healthcare ecosystem.
28. **Investment strategy for the VHT:** to ensure the successful development and implementation of the VHT, continuous, robust and diverse sources of support, both public and private, are essential.

This support will facilitate a wide range of activities from existing platform maintenance, over the integration of new models and modalities, to ensuring legal compliance. A long-term investment strategy by the EU, including new EU calls for proposals, will attract multi-funder contributions and encourage co-investment from Member States and private partners in developing the VHT infrastructure and applications. Moreover, securing long-term funded staff and promoting resource reuse will further contribute to the sustainability and advancement of the VHT ecosystem.

29. **Integrating the VHT into the European infrastructure landscape:** European policymakers should champion policy innovation to accelerate VHT development and adoption. The VHT's ambitious scope requires its integration into the European Infrastructure landscape. A decentralized approach is recommended, establishing national nodes to promote collaboration and resource sharing across Europe (*cfr.* EOSC). Establishing the VHT as a European Digital Infrastructure Consortium or forming a dedicated European Partnership for the VHT will provide a sustainable framework for its long-term development, implementation and ecosystem engagement. This necessitates the active involvement of EU Member States as well as public and private partners to pool resources, expertise, and infrastructure, fostering collaboration across national borders.
30. **The Importance of environmental sustainability for the VHT:** the development and deployment of the VHT needs to include careful consideration of its environmental impact. Minimising the environmental footprint of the VHT infrastructure requires examining the energy consumption associated with data storage, processing, and model execution. This necessitates adopting sustainable practices in data centre operations and exploring energy-efficient algorithms and hardware architectures. Embracing the principles of Planetary Health, the VHT should strive to contribute to a healthier planet, recognizing the interconnectedness between human health and the environment. This underscores the importance of applying a "sustainability triangle" approach to VHT development, encompassing environmental, social, and economic considerations to ensure its long-term viability.

35 Recommendations for the stakeholders of the VHT ecosystem

This chapter outlines recommendations for successfully rolling out the Virtual Human Twin (VHT) over the next decade, focusing on synergies with and dependencies on existing initiatives in research, legislation, and regulatory landscapes. These recommendations are categorized by stakeholder groups to identify areas in which each of them can contribute to the successful development and roll out of the VHT.

35.1 European Commission

- **Establish dedicated funding and deployment programs:** dedicate funding specifically for advancing VHT research, development and deployment. This includes supporting basic research on identified knowledge gaps, translational research efforts, and projects with varying Technology Readiness Levels (TRLs). Funding should cover both generic/population-specific and personalized Digital Twins. Prioritize research and innovation in developing, testing, validating, and verifying advanced VHT technologies while ensuring synergy with existing digital services and capabilities at the European level. Create a European Partnership for VHT.
- **Champion a European VHT initiative:** continue to champion the further development and deployment of a comprehensive European VHT initiative to accelerate the best use of technologies in healthcare, taking advantage of existing European VHT champions and assets. This includes fostering a deeper understanding of AI-based and simulation tools among healthcare providers, clinicians, as well as an understanding of clinical and patient needs among VHT researchers and developers. The European Commission's initiative involves procuring a Platform for Advanced Virtual Human Twin Models, showcasing its commitment to funding, training and new research and innovation in this crucial area.
- **Promote interaction with existing EU-level initiatives:** ensure the VHT initiative aligns with other relevant European initiatives like the European Health Data Space and others to leverage existing infrastructure and facilitate data sharing. Support federated infrastructures in digital health and research enhancing interoperability and building of VHT.
- **Support and take decision on the European Digital Infrastructure (EDIC) for the VHT:** participate on the co-design and planning on the formation of a new EDIC to support the VHT, with participation of at least three Member States.

35.2 EU Member states

- **Support National and Regional VHT Initiatives:** encourage and fund regional and national initiatives, including dedicated centres, grassroots projects, and funding programs, that promote the development of Digital Twins in healthcare. This should include mapping and refining these initiatives at the member state level to foster collaboration and optimise resource allocation.
- **Communicate and disseminate funding options:** create awareness on the possible funding pathways so that the VHT ecosystem can understand when and how they can leverage the different funding options, according to each country's strategic objective.
- **Foster Collaboration with the European VHT Platform:** align national healthcare strategies with the development of the European VHT platform, facilitating data sharing and promoting the adoption of VHT technologies within national healthcare systems. Co-establish the European Partnership for VHT.
- **Engage in Federated Governance:** Open the discussion on federated governance to harmonise and facilitate cross-country regulations and VHT use.

35.3 Research community

- **Address fundamental research challenges:** focus on research to bridge foundational gaps in statistics, mathematics, and computing related to Digital Twins. This involves conducting rigorous research on model development, validation, and verification, as well as advancing methods for data generation, including, but not limited to, real-time simulations, multi-scale modelling, uncertainty quantification, and decision-making using Digital Twins.

- **Develop and share VHT resources:** contribute to building the VHT knowledge base by developing, validating, and sharing data, models, algorithms, and good practices. This can be achieved by actively contributing resources to the VHT infrastructure.
- **Address unmet stakeholder needs:** engage with all stakeholders to develop fit-for-purpose Digital Twins in a co-creation process, aligned with end-user needs and workflows. Implement high-impact validated use cases in clinical practice, prioritizing low hanging fruits with clinical champions to foster broader end-user engagement and uptake.
- **Integration with broader EU4Health objectives:** the VHT research field should align with EU4Health objectives, such as enhancing healthcare resilience, promoting equitable access to healthcare, and fostering the digital transformation of health systems. Integration with these broader goals will ensure the VHT contributes to a unified and impactful health strategy across the EU.

35.4 Industry

- **Engage in public-private partnerships:** collaborate with research institutions and public entities to accelerate VHT development and translation into practical healthcare solutions.
- **Explore VHT-based business models and VHT intermediary businesses:** identify and develop viable business models around VHT technologies, such as providing VHT-as-a-Service, developing VHT-powered software solutions for healthcare providers, creating personalized healthcare applications for individuals or taking part with intermediary businesses. This can involve collaborating with other stakeholders to establish a VHT marketplace where resources and services can be exchanged.
- **Contribute to standardization efforts:** actively participate in defining and implementing standards for VHT data, models, and workflows to ensure interoperability and facilitate wider adoption of VHT technologies.

35.5 Healthcare providers

- **Engage in collaborative research:** participate actively in collaborations with VHT developers to provide clinical insights, define priorities, and guide the design of solutions that address real-world patient needs and integrate seamlessly into existing workflows.
- **Advocate for training and tools:** advocate for and participate in the development of training programs and user-friendly tools that simplify the interpretation of VHT-generated data, ensuring healthcare professionals are equipped to utilize the technology effectively in clinical practice.
- **Contribute to validation efforts:** support and contribute to rigorous clinical trials and validation studies to establish the impact of VHT on patient outcomes, cost-efficiency, and clinical processes, while addressing concerns about data security, algorithmic bias, and patient privacy.

35.6 Regulatory bodies and standards organisations

- **Develop VHT-specific regulatory pathways:** establish clear regulatory guidelines and pathways for the development, validation, and clinical use of VHT applications, ensuring patient safety and data privacy. This might involve establishing sandboxes for testing and evaluating VHT applications in a controlled environment.
- **Develop VHT-specific standards:** clear definitions for VHT software components and uniform standards for data formats, ontologies, and workflows are essential to streamline approval processes, enhance interoperability, and support reproducibility.
- **Collaborate with international organizations:** Participate in international efforts to harmonise regulatory approaches to VHT, promoting a globally consistent and robust framework for its development and use.

35.7 Payer, Buyer, and Health Technology Assessment Community

- **Utilise innovation adoption instruments:** HTA bodies and payers should leverage tools like pre-commercial procurement and value-based procurement to generate evidence, evaluate cost structures, and establish reimbursement pathways for VHT technologies.
- **Standardise evaluation and payment models:** Clear criteria for assessing clinical and economic value should be developed alongside innovative payment mechanisms, such as value-based agreements, to address long-term uncertainties and support adoption.
- **Ensure equitable access and harmonization:** Policies must promote affordability, accessibility, and harmonized evaluation strategies across Member States to unlock the full potential of VHT and improve healthcare outcomes.

35.8 Ethical, legal, and social implication experts

- **Conduct Comprehensive Assessment of Ethics and Social Implications:** Continuously assess the ethical implications of VHT development and deployment, focusing on issues such as data privacy, informed consent, human control, bias in algorithms, and potential societal impacts.
- **Develop Ethical Guidelines and Codes of Conduct:** Develop and promote ethical guidelines and codes of conduct for all stakeholders involved in VHT research, development, and application. This includes addressing the ethical and legal aspects of reusing and integrating data for VHT development.
- **Engage in Public Dialogue:** Facilitate public dialogue on the ethical and societal implications of VHT to ensure societal needs, values and concerns are carefully considered, transparency is promoted, and trust relationships are established and maintained.

35.9 Patients and the public

- **Active Participation in VHT Development:** encourage patients and the public to actively participate in shaping the VHT by providing feedback on its development, raising concerns, and contributing to discussions on ethical, legal and social issues. Patient perspectives are crucial in ensuring that the VHT is developed and used in a way that aligns with their values and needs.
- **Education and Awareness:** promote education and awareness programs to inform the public about the potential benefits, risks, and ethical considerations associated with VHT. Use immersive technologies coupled to VHT technologies to lower the barriers for access and understanding.

The successful realisation and roll-out of the VHT hinge on a collaborative and coordinated approach from all stakeholder groups. These recommendations provide a starting point for addressing the technical, infrastructural, ethical, legal, and societal challenges associated with the VHT's development, ensuring its responsible integration into the future of healthcare.

36 Annex 1: contributors

Consortium			
First name	Surname	E-mail	Country
Jazmin	Aguado-Sierra	Barcelona Supercomputing Center	Spain
Katrin	Amunts	Forschungszentrum Jülich	Germany
Spiros	Athanasiou	Athena Research Center	Greece
Sophie	Bekisz	University of Liege	Belgium
Laura	Bevis	Queen Mary University of London	UK
Charles	Boulitrop	Inria	France
Sebastien	Bratieres	Pi School	Italy
Beatrice	Bressan	Lynkeus	Italy
Marian	Bubak	ACC Cyfronet AGH Krakow	Poland
Desirée	Caccaro	VPH Institute	Belgium
José	Carbonell Caballero	Barcelona Supercomputing Centre	Spain
Nathan	Carvalho	DIGITALEUROPE	Belgium
Francesca	Conte	University of Bologna	Italy
Martina	Contin	VPH Institute	Italy
Marilisa	Cortesi	University of Bologna	Italy
Lorenzo	Cristofaro	Lynkeus	Italy
Cristina	Curreli	University of Bologna	Italy
Giorgio	Davico	University of Bologna	Italy
Giacomo	de Feo	Lynkeus	Italy
Mirko	De Maldè	Lynkeus	Italy
Roberta	De Michele	VPH Institute	Italy
Lotte	De Schrijver	École Polytechnique Fédérale de Lausanne	Switzerland
Simon	Denil	VITO	Belgium
Thomas	Desaive	University of Liege	Belgium
Timo	Dickscheid	Forschungszentrum Jülich	Germany
Dirk	Drasdo	Inria	France
Ludovica	Durst	Lynkeus	Italy
Bart	Elen	VITO	Belgium
Gokhan	Ertaylan	VITO	Belgium
Enzo	Fabiani	Pi School	Italy
Francesco	Fedele	Pi School	Italy
Barry	Fitzgerald	VPH Institute	Belgium
Dimitrios	Fotiadis	FORTH	Greece
Minos	Garofalakis	Athena Research Center	Greece
Liesbet	Geris	VPH Institute	Belgium
Elfi	Goesaert	VITO	Belgium
Martin	Golebiewski	Heidelberg Institute for Theoretical Studies	Germany
Björn	Hagemeier	Forschungszentrum Jülich	Germany
Alfons	Hoekstra	Universiteit van Amsterdam	The Netherlands
Valentina	Huber	Forschungszentrum Jülich	Germany
Yannis	Ioannidis	Athena Research Center	Greece
Adrian	Ionescu	École Polytechnique Fédérale de Lausanne	Switzerland
Alba	Jene	Barcelona Supercomputing Centre	Spain

Marc	Joiret	University of Liege	Belgium
Frederic	Jung	VITO	Belgium
Joyce	Kao	Uniklinik RWTH Aachen	Germany
Georgia	Karanasiou	FORTH	Greece
Sofia	Karvounari	Athena Research Center	Greece
Marek	Kasztelnik	ACC Cyfronet AGH Krakow	Poland
Aaron	Lee	Queen Mary University of London	UK
Raphaëlle	Lesage	VPH Institute	Belgium
Marco	Liotta	University of Bologna	Italy
Helena	Lovegrove	DIGITALEUROPE	Belgium
Luiz Carlos	Maia Ladeira	University of Liege	Belgium
Evita	Mailli	Athena Research Center	Greece
Maciej	Malawski	ACC Cyfronet AGH Krakow	Poland
Daniel	Mallmann	Forschungszentrum Jülich	Germany
Flaminia	Malvezzi	Lynkeus	Italy
Daniella	Manassero	DIGITALEUROPE	Belgium
Jose	Martinez-Usero	DIGITALEUROPE	Belgium
Gernot	Marx	Uniklinik RWTH Aachen	Germany
Anthony	Mathur	Queen Mary University of London	UK
Gerhard	Mayer	Heidelberg Institute for Theoretical Studies	Germany
Jan	Meizner	ACC Cyfronet AGH Krakow	Poland
Sabato	Mellone	University of Bologna	Italy
Massimiliano	Mercuri	University of Bologna	Italy
Arnaud	Montagud	Barcelona Supercomputing Centre	Spain
Davide	Montesarchio	VPH Institute	Italy
Hannah	Morley Fletcher	Lynkeus	Italy
Serena	Moscato	University of Bologna	Italy
Wolfgang	Mueller	Heidelberg Institute for Theoretical Studies	Germany
Filippo	Nardini	University of Bologna	Italy
Pjotr	Nowakowski	ACC Cyfronet AGH Krakow	Poland
Sanjay	Pant	Lynkeus	Italy
Giulia	Pederzani	University of Amsterdam	The Netherlands
Steffen	Petersen	Queen Mary University of London	UK
Artem	Platonov	VPH Institute	Belgium
Richard	Polzin	Uniklinik RWTH Aachen	Germany
Alexandra	Prodan	Empirica	Germany
Janaki	Raman Rangarajan	VPH Institute	Belgium
Amaryllis	Raouzaïou	Athena Research Center	Greece
Matilde	Ratti	University of Bologna	Italy
Elisa	Rauseo	Queen Mary University of London	UK
Gaia	Reale	Lynkeus	Italy
Alice	Reid	Queen Mary University of London	UK
Caroline	Roney	Queen Mary University of London	UK
Bernd	Schuller	Forschungszentrum Jülich	Germany
Andreas	Schuppert	Uniklinik RWTH Aachen	Germany
Maxime	Semersant	Inria	France

Ashwani	Sharma	VPH Institute	Belgium
Greg	Slabaugh	Queen Mary University of London	UK
Stan	Snowball	VPH Institute	Belgium
Yann	Sprunger	École Polytechnique Fédérale de Lausanne / Xsensio	Switzerland
Goran	Stanic	VPH Institute	Belgium
Bernard	Staumont	University of Liege	Belgium
Veli	Stroetmann	Empirica	Germany
Michael	Strubin	DIGITALEUROPE	Belgium
Michael	Tarnawa	Forschungszentrum Jülich	Germany
Rainer	Thiel	Empirica	Germany
Diem	Tram	DIGITALEUROPE	Belgium
Antonella	Trezzani	Lynkeus	Italy
Vincent	Uyttendale	University of Liege / Insilicare	Belgium
Alfonso	Valencia	Barcelona Supercomputing Centre	Spain
Zita	Van Horenbeeck	VPH Institute	Belgium
Lex	van Loon	University of Amsterdam	The Netherlands
Mariano	Vasquez	Barcelona Supercomputing Centre	Spain
Marco	Viceconti	University of Bologna	Italy
Irene	Vignon-Clementel	Inria	France
Malte	von Tottleben	Empirica	Germany
Gabor	Závodszky	University of Amsterdam	The Netherlands

Advisory Group of Stakeholders

First name	Surname	E-mail	Country
Peter	Aulbach	Siemens Healthineers	Germany
Maria	Benedetti	University of Bologna	Italy
Mikael	Benson	Karolinska Institutet	Sweden
Patrick	Boisseau	MedTech Europe	Belgium
Fidelia	Cascini	Università Cattolica S Cuore	Italy
Paola	Grosso	University of Amsterdam	The Netherlands
Ger	Janssen	Philips	The Netherlands
Robert	Madjono	TüvSüd	Germany
Gernot	Marx		
Signe	Mezinska	University of Latvia	Latvia
Julia	Neguer	Dassault Systèmes	France
Julio	Saez-Rodrigo	Heidelberg University	Germany
Annabel	Seebohm	COCIR Aisbl	Belgium
Maxime	Sermensant		
Valentina	Strammiello	European Patients' Forum	Belgium
Mariano	Vasquez		

Industry Advisory Board

First name	Surname	Affiliation	Country
Paymam	Afshari	DePuy Synthes Spine, Johnson & Johnson	USA
Sophie	Batas	Dassault Systèmes	France
Vincenzo	Carbone	InSilicoTrials Technologies	Italy
Jean	Colombel	Dassault Systèmes	France
Martha	De Cunha	Edwards Lifesciences	The Netherlands
Luca	Emili	InSilicoTrials Technologies SPA	Italy

Oliver	Frings	Siemens Healthineers	Germany
Malte	Fritsche	Bitkom	Germany
Elisa	Gastaldi	Siemens Healthineers	Belgium
Gordon	Johnston	Johnson & Johnson	USA
Maria Cristina	Jori	Mediolanum Cardio Research	Italy
Lance	Ladic	Siemens Healthineers	USA
Thierry	Marchal	ANSYS	Belgium
Shiny	Martis	Nova In Silico / VCLS	France
Chris	Morton	ELEM	Spain
Antoine	Perrier	Twinsight Medical	France
Emmanuel	Pham	Nova In Silico	France
Melanie	Poulin-Costello	Roche	Switzerland
Cecile	Rousseau	VCLS	France
Simon	Sonntag	Virtonomy GmbH	Germany
Visa	Suomi	Mathworks	Germany
Olaf	Van Der Sluis	Philips	Netherlands
Danny	Van Roijen	Agoria	Belgium
Emmanuelle	Voisin	VCLS	France
Contributors			
First name	Surname	Affiliation	Country
Michiel	Adriaens	Maastricht University	The Netherlands
Ankush	Aggarwal	University of Glasgow	UK
Sahar	Aghakhani	Université d'Évry	France
Mariano	Alcañiz	Universidad Politécnica de Valencia - Spain	Spain
Alessandra	Aldieri	Politecnico Torino	Italy
Rossana	Alessandrello	AQuAS – Agency for Health Quality and Assessment of Catalonia	Spain
Nada	Alkhayat	European Commission	Belgium
Gary	An	University of Vermont	USA
Odysseos	Andreani	Epos iasis	Cyprus
Antonio	Andreu	Eatris Eric	The Netherlands
Koji	Andriamahery	Stella Surgical / Mines Ales	France
Juan	Arenas Marquez	ELIXIR Europe	UK
Finbar	Argus	University of Auckland	New Zealand
Andreas	Arndt	BIOTRONIK SE & Co. KG	Germany
Andreas	Arnegger	OSORA medical GmbH	Germany
Kartik	Arora	Siemens Healthineers	Germany
Ilja	Arts	Maastricht University	Germany
Muhammad	Aslani	Radboud University	The Netherlands
Roger	Assaker	MDsim	Luxembourg
Folkert	Asselbergs	Amsterdam UMC & University of Amsterdam	The Netherlands
Koray	Atalag	GALATA-Digital	Turkey
Kevin	Aubert	IsiTwin	France
Franck	Auge	Sanofi	France
Luca	Azzolin	Numericor	Austria
Thiranja Prasad	Babarenda Gamage	University of Auckland	New Zealand

Aldo	Badano	US FDA	USA
Samo	Badr	Logex Group	The Netherlands
Steven	Bagshaw	CPI	UK
Hasan	Balci	Maastricht University	The Netherlands
Annabelle	Ballesta	Inserm	France
Rudi	Balling	University of Bonn	Germany
Pieter	Ballon	VUB (Free University Brussels)	Belgium
Prith	Banerjee	ANSYS	USA
Michele	Barbier	Inria	France
manuele	barbieri	Rina	Italy
Alessia	Baretta	InSilicoTrials Technologies SPA	Italy
Emmanuel	Barillot	Institut Curie	France
Tom	Battisti	Dassault Systèmes	France
Erica	Beaucage-Gauvreau	KU Leuven	Belgium
Elisabeth	Beck	DIN - German Institute for Standardization	Germany
Mirza Faisal	Beg	Voronoi Health Analytics Incorporated	Canada
Jeroen	Belien	Amsterdam UMC	The Netherlands
Ivan	Benemerito	Sheffield University	UK
Sonja	Bergner	Eurice	Germany
Job	Berkhout	National Institute for Public Health and the Environment	The Netherlands
Francesca	Berti	Politecnico di Milano	Italy
Cristobal	Bertoglio	University of Groningen	The Netherlands
Olivier	Beuf	CREATIS CNRS 5220, Inserm	France
Fanny	Beynaerts	MSD	France
Pinaki	Bhattacharya	University of Sheffield	UK
Elisabetta	Biasin	KU Leuven	Belgium
Jonas	Biehler	Ebenbuild	Germany
Florence	Bietrix	EATRIS	The Netherlands
Claire	Biot	Dassault Systèmes	France
Annette	Birkhold	Siemens Healthineers	Germany
Niklas	Blomberg	IHI JU	Belgium
Leone	Bock	Lygature	The Netherlands
Francois	Boissel	Nova In Silico SA	France
Guido	Bologna	University of Applied Sciences and Arts of Western Switzerland	Switzerland
Cecilia	Bonefeld-Dahl	DIGITALEUROPE	Belgium
Cecile	Bonnard	Dassault Systèmes	France
Luca	Bontempi	Erasmus MC	The Netherlands
Lucas	Böttcher	Frankfurt School of Finance & Management	Germany
Redouane	Boudjemaa	Blida University 1	Algeria
Steven	Bourke	Personal Pulse	Switzerland
Vincent	Bouwman	4RealSim Services BV	The Netherlands
Ramzi	Bouzaïenne	ENIT	Tunisia
Thibault	Boyer	École Supérieure d'Ingénieurs Léonard de Vinci	France
Carson	Bradbury	Siemens	UK

Sara	Bridio	European Commission, Joint Research Center, JRC.F	Italy
Ciaran	Briscoe	Geneera	Ireland
Oscar	Brück	Helsinki University Hospital	Finland
Jan	Brüning	Charité - Universitätsmedizin Berlin	Germany
Lutz	Brusch	TU Dresden	Germany
Rebecca	Bryan	Synopsys	USA
Roland-Valentin	Bumbuc	Amsterdam UMC	The Netherlands
Carmen	Buzea	Transilvania University of Brasov	Romania
Alice	Byram	RSM Digital Health Section	UK
Oscar	Camara	Universitat Pompeu Fabra	Spain
Jenny	Camaradou	Health clusterNet	UK
Claudio	Capelli	University College London	UK
Antonella	Carbonaro	University of Bologna	Italy
Chiara	Carboni	TU Dresden	Germany
Aurélie	Carlier	Maastricht University	The Netherlands
Svenja	Caspers	Research Centre Jülich	Germany
Gastone	Castellani	IRCCS AOUBO	Italy
Filippo	Castiglione	Consiglio Nazionale delle Ricerche (CNR)	Italy
Gunnar	Cedersund	Linköping University	Sweden
Florian	Centler	University of Siegen	Germany
Mario	Ceresa	European Commission, Joint Research Center, JRC.F	Belgium
Ledoux	Charles	ETH Zurich	Switzerland
Michel	Chaudron	Eindhoven University of Technology	The Netherlands
Sara	Checa	Charité -Universitätsmedizin Berlin	Germany
Chen	Chen	University of Sheffield	UK
Lorenzo	Chiari	University of Bologna	Italy
Davide	Chieco	Computational Life Inc.	Spain
Julie	Choisne	University of Auckland	New Zealand
Ioanna	Chouvarda	Aristotle University of Thessaloniki	Greece
Georges	Christé	Université Lyon1	France
Catherine	Chronaki	HL7 Foundation	France
Patrick	Clarysse	CREATIS CNRS 5220, Inserm	France
Patrick	Clarysse	CREATIS CNRS 5220, Inserm	France
Piet	Claus	KU Leuven	Belgium
Laure-Alix	Clerbaux	European Commission	Belgium
Andreu	Climent	Universitat Politècnica de Valencia	Spain
Rachel	Clipp	Kitware	USA
Matthijs	Cluitmans	CARIM	The Netherlands
Robert	Cockrell	University of Vermont	USA
Mitchel	Colebank	University of California Irvine	USA
Maxime	Colomb	Inria	France
Filippo	Colombo	Mediolanum Cardio Research Srl	Italy
David	Constant	Happycity	Canada
Christian	Contarino	Computational Life Inc.	Spain
Christophe	Cop	Konsolidate	Belgium
Sylvie	Couderc	BD	France

Guy	Courbebaisse	CREATIS CNRS 5220, Inserm	France
Alberto	Credi	University of Bologna	Italy
Milena	Crespi	AIMED trial srl	Italy
Luca	Cristofolini	University of Bologna	Italy
Enrico	Dall'Ara	Sheffield University	UK
Martin	Daumer	SLC e.V. - The Human Motion Institute	USA
Elizabeth	De Lange	Leiden University	The Netherlands
Karen	De Leener	Materialise	Belgium
Laia	de Nadal	Universitat Pompeu Fabra	Spain
Maarten	De Vos	KU Leuven	Belgium
Lieselot	Declercq	Ghent University	Belgium
Marios	Demetriades	University of Cyprus	Cyprus
Oleg	Demin Jr	InSysBio	Cyprus
Charles	Desrosiers-Gagnon	Université de Montréal	Canada
Günther	Deuschl	University of Kiel	Germany
Jan	D'hooge	KU Leuven	Belgium
Yanis	Diallo	ESILV	France
Francisco	Diaz-Mitoma	Bowhead Health	Canada
Mauro	DiNuzzo	Netabolics SRL	Italy
Hubert	Dirven	Norwegian Institute of Public Health	Norway
Sagar	Dolas	SURF	The Netherlands
Joaquin	Dopazo	Junta De Andalucia	Spain
Bernadette	Dorizzi	Telecom SudParis	France
Noemie	Dubruel	Université Toulouse 1	France
Tino	Ebbers	Linköping University	Sweden
Moein	Einollahzadeh Samadi	Uniklinki RWTH Aachen University	Germany
Seyed Ali	Elahi	KU Leuven	Belgium
Wahbi	El-Bouri	University of Liverpool	UK
Elisa	Elhadj	KU Leuven	Belgium
Noha	Elnahas	Independent	Belgium
Imane	Elonen	Turku University of Applied Sciences	Finland
Carl-Leandro	Enslin	Karlsruhe Institute of Technology	Germany
Marina	Esteban	ETH Zurich	Switzerland
Christopher	Esterhuyse	University of Amsterdam	The Netherlands
Ewan	Evain	Dassault Systèmes	France
Marco	Evangelos Biancolini	Università di Roma Tor Vergata	Italy
Hana	Fadlallah	SanteCie	France
Francois	Fages	Inria	France
Baylor	Fain	Texas Christian University	USA
Nele	Famaey	KU Leuven	Belgium
Margherita	Fanos	European Commission	Luxembourg
Anne-charlotte	Fauvel	EATRIS	The Netherlands
Emmanuel	Federbusch	IROC	France
Heleen	Fehervary	KU Leuven	Belgium
Kyriacos	Felekkis	University of Nicosia	Cyprus
Behrooz	Fereidoon nezahad	TU Delft	The Netherlands
Sandra	Ferreira	Universidade Beira Interior	Portugal

Jan	Filakovsky	BIOMAPAs	Lithuania
Milena	Fini	Rizzoli Orthopaedic Institute	Italy
Jasmin	Fisher	University College London	UK
Asmund	Flobak	Norwegian University of Science and Technology	Norway
David	Flynn	Boston Scientific	USA
Luis	Fonseca	University of Florida	USA
Nicolo	Formento	École Supérieure d'Ingénieurs Léonard de Vinci	France
Daniel	Fougeres	Association MICADO	France
Eriberto Andrea	Franchi	Politecnico di Milano	Italy
Alejandro	Frangi	University of Manchester	UK
Bertrand	Fréchède	Université Claude Bernard Lyon 1 & Université Gustave Eiffel	France
Monika	Frenzel	French National Research Agency	France
Fabian	Froehlich	Harvard Medical School	USA
Holger	Fröhlich	Fraunhofer Gesellschaft e.V.	Germany
Morgane	Garreau	Inria	France
Tatiana	Gassmann	Karlsruhe Institute of Technology	Germany
Nicolas	Gaudenzio	Inserm	France
Jean-Frédéric	Gerbeau	National Institute for Research in Digital Science and Technology	France
Philippe	Gesnouin	Inria	France
Hamza	Gharafi	Arias Tech Solutions	Morocco
Frank	Gijzen	Erasmus MC	The Netherlands
Kristina	Giske	German Cancer Research Center	Germany
William	Glascoe	Robert Morris University	USA
James	Glazier	Indiana University	USA
Carole	Globe	The University of Manchester	UK
Veronique	Gobry	Sciprom	Switzerland
Philipp	Goedecker	COCIR Aisbl	Belgium
Marleen	Golsteijn - Odio Moret	VU Amsterdam	The Netherlands
Lieselot	Gorter	Utrecht University	The Netherlands
Yaneck	Gottesman	Telecom SudParis	France
Nils	Götzen	4RealSim Services BV	The Netherlands
Leonid	Goubergrits	Charité - Universitätsmedizin Berlin	Germany
Emeline	Gougeon	Comité scientifique Pro Anima	France
Vittorio	Grieco	University of Catania	Italy
Corina	Guder	University of Bergen	Norway
Jose M.	Guerrero	InfoSeg	Spain
Maria	Guillem	Universitat Politecnica de Valencia	Spain
Bhumika	Gupta	Open Health Systems Laboratory	USA
Stefan	Haar	Inria	France
Jens	Habermann	BBMRI-ERIC	Austria
Margaret	Hagen	University of Bergen	Norway
Kevin	Hakkakian	Inria	France
Bernhard	Hametner	Austrian Institute of Technology	Austria

Ira	Haraldsen	Oslo University Hospital	Norway
Vinod Pradyumna	Harlapur	Indian Institute of Science	India
Tilo	Hartmann	VU Amsterdam	The Netherlands
Houda	Hassini	Telecom SudParis	France
Kyriacos	Hatzaras	European Commission	Luxembourg
Alexandra	Hauguel	Hôpital Marie Lannelongue, Ladhyx Ecole Polytechnique	France
Brian	Helba	Kitware	USA
Tomas	Helikar	University of Nebraska–Lincoln	USA
Leif	Hellevik	Norwegian University of Science and Technology	Norway
Ahmed	Hemedan	University of Luxembourg	Luxembourg
Rick	Hendriks	Hogeschool Leiden	The Netherlands
Ulf	Hengstmann	Bayer	Germany
Esteban	Hernandez	University of Idaho	USA
Robert	Hester	HC Simulation, LLC	USA
James	Hetherington	University College London	UK
Donald	Hickey	Pumpinheart	Ireland
Chi	Hieu Le	University of Greenwich	UK
Rafiazka	Hilman	University of Amsterdam	The Netherlands
Harvey	Ho	University of Auckland	New Zealand
Stefab	Höhme	Leipzig University	Germany
YongKai	Hong	Dalian University of Technology	China
Wouter	Huberts	Eindhoven University of Technology	The Netherlands
Miguel	Hueso	Hospital Bellvitge	Spain
Eystein	Husebye	Bergen Hospital Trust	Norway
Jari	Hyttinen	Tampere Universities	Finland
Yaron	Ilan	Hebrew University	Israel
Matteo	Italia	Politecnico di Milano	Italy
Thomas	Ittel	Uniklinik RWTH Aachen	Germany
Lucian	Itu	Transilvania University of Brasov	Romania
Tatjana	Ivaskiene	imcentras	Italy
Ivar	Janmaat	SURF	The Netherlands
Jorg	Janssen	Lygature	The Netherlands
Luis David	Jimenez	Esqlabs	Germany
Liuyi	Jin	Texas A&M University	USA
Viktor	Jirsa	Aix Marseille University	France
John	Jobin	Chalmers University	Sweden
Anna	Jones	University of Liverpool	UK
Ilse	Jonkers	KU Leuven	Belgium
Lumens	Joost	Maastricht University	The Netherlands
Auste	Kanapeckaite	multiple	Lithuania
Michele	Kanhonou	ESILV	France
Panagiotis	Karampelesis	University of Patras	Greece
Walter	Karlen	Ulm University	Germany
Matts	Karlsson	Linköping University	Sweden
Koen	Kas	Healthskouts	Belgium
Elvan	Kaya	IBM	USA

Sarah	Keating	University College London	UK
Mustafa	Khammash	ETH Zürich	Switzerland
Sanjay	Kharche	Victoria Hospital	Canada
Takoua	Khorchani	ECRIN	Paris
Wouter	Klijn	Forschungszentrum juelich	Germany
Marina	Klitsi	Athens Technology Center	Greece
Adam	Knapp	University Of Florida	USA
Carole	Knibbe	Inria	France
Adam	Knowles	Liverpool John Moores University	UK
Rüdiger	Köhling	University of Rostock	Germany
Walter	Kolch	University College Dublin	Ireland
Miroslav	Koncar	Philips	The Netherlands
Matthias	Konig	Humboldt University Berlin	Germany
Marko	Korenjak	European Liver Patients' Association	Belgium
Bertrand	Kornfeld	PDCApIy	France
Peter	Kottman	Inria	France
Iordanis	Koutsopoulos	Athens University of Economics and Business	Greece
Dennis	Koverzin	Life Elevated Inc.	Canada
Jérôme	Kowalski	Inria	France
Oliver	Kraft	Karlsruhe Institute of Technology	Germany
Valeria	Krzyszhanovskaya	University of Amsterdam	The Netherlands
Alexander	Kulesza	Nova In Silico	France
Tomáš	Kulháněk	VITO	Belgium
Maneesh	Kumar	Cardiff University	UK
Luigi	LaBarbera	Politecnico di Milano	Italy
Damien	Lacroix	University of Sheffield	UK
Yoann	Lafon	Université Lyon, Université Gustave Eiffel	France
Jeremy	Laforet	BMBI UMR7338 CNRS	France
Juntong	Lai	University of Sheffield	UK
Colin	Lake	UCB	Belgium
Pablo	Lamata	King's College London	UK
Micha	Landoll	RWTH Aachen University	Germany
Konrad	Lang	BBMRI-ERIC	Austria
Christophe	Lanneau	Genopole	France
Briane	Laruy	TüvSüd	Germany
Mojtaba	Lashgari	University of Oxford	UK
Reinhard	Laubenbacher	University of Florida	USA
Calzone	Laurence	Institut Curie	France
Miroslav	Lávička	University of West Bohemia	Czech republic
Job	Leenen	Isala Hospital	The Netherlands
Janne	Lehtio	Karolinska Institutet	Sweden
Ignace	Lemahieu	Ghent University	Belgium
Silvia	Lenaerts	Eindhoven University of Technology	The Netherlands
Gabriele	Leoni	European Commission, Joint Research Center, JRC.F	Italy
Raphaëlle	Lesage	Esq Labs	Germany
Steven	Levine	Dassault Systèmes	France

Maria Luisa	Levitano	Università Milano Bicocca	Italy
Xinshan	Li	University of Sheffield	UK
Elisa	Lievevrouw	KU Leuven	Belgium
Chinchien	Lin	University of Auckland	New Zealand
Pietro	Lio	University of Cambridge	UK
Walter	Lioen	SURF	The Netherlands
Bao	Liu	Fudan University	China
Raffaele	Lodi	University of Bologna	Italy
Josh	Loecker	University of Nebraska–Lincoln	USA
Axel	Loewe	Karlsruhe Institute of Technology	Germany
Luis	Lopera	Friedrich-Alexander Universität Erlangen-Nürnberg	Germany
Paulo	Lopes	National Education and Research Network (RNP)	Brasil
Nicola Francesco	Lopomo	University of Brescia	Italy
Matthias	Louws	University of Amsterdam	The Netherlands
Daniel	Lundqvist	Karolinska Institutet	Sweden
Giulia	Luraghi	Politecnico di Milano	Italy
Andreas	Lysandropoulos	Parexel	Belgium
Rob	MacLeod	University of Utah	USA
Alejandro	Madrid Valiente	UPF	Spain
Shivani Kaushal	Maffi	Open Health Systems Laboratory	USA
Florian	Maier	Simq GmbH	Germany
Alessandro	Maiocchi	Bracco spa	Italy
Marco	Manca	Scimpulse	The Netherlands
Lorenzo	Manganaro	aizoon	Italy
Sabrina	Mann	DIN - German Institute for Standardization	Germany
Davide	Marche		Italy
Francesca	Margara	University of Oxford	UK
Sergio	Marín	IRB Barcelona	Spain
Michele	Marino	Università degli Studi di Roma Tor Vergata	Italy
Henk	Marquering	Amsterdam University Medical Centers	The Netherlands
Sergi	Marsol Torrent	Barcelona Supercomputing Center	Spain
Juan Pablo	Martínez	University of Zaragoza	Spain
Dorel	Mathurin	Max Plank Institute for Molecular Genetics	Germany
Lisa	Matthews	NYU Grossman School of Medicine	USA
Giancarlo	Mauri	Università di Milano-Bicocca	Italy
Jon	Mc Cullough	University College London	UK
Eadaoin	Mc Kiernan	University College Dublin	Ireland
Jeannine	McCune	City of Hope	USA
Borna	Mehrad	University of Florida	USA
Roel	Meiburg	Inria	France
Jorge	Mejias	University of Amsterdam	The Netherlands
Nedra	Mellouli	École Supérieur d'Ingénieurs Léonard de Vinci	France
Marlene	Mengoni	University of Leeds	UK

Francesco	Messina	INMI Lazzaro Spallanzani National Institute of Infectious Diseases	Italy
Miriam	Meyer	Siemens Healthineers	Germany
Christian	Michael	University of Michigan	USA
Heleen	Middelkamp	EUROoCS	The Netherlands
Francesco	Migliavacca	Politecnico di Milano	Italy
Gary	Mirams	University of Nottingham	UK
David	Mitton	Université Claude Bernard Lyon 1,	France
Sabyasachi	Mohanty	University of Nebraska–Lincoln	USA
Ikram	Mohout	KU Leuven	Belgium
Philippe	Moingeon	Servier R&D	France
Mahsa	Momtahan	Movv Robotics	Belgium
Benoit	Monegier	IP Paris	France
Pedro	Monteiro	Universidade de Lisboa	Portugal
Peter	Moon	VIRTACH	UK
Hernan	Morales	Dassault Systèmes	France
Enrique	Morales Orcajo	Ambu	Spain
Heike	Moser	DIN - German Institute for Standardization	Germany
Omer	Mujahid	University of Girona	Spain
Lucas	Muller	University of Trento	Italy
Vivek Sheraton	Muniraj	University of Amsterdam	The Netherlands
Amalia	Munoz Pineiro	European Commission, Joint Research Center, JRC.F	Belgium
Marta	Musidlowska	KU Leuven	Belgium
Arun	Nadarasa	International Social Prescribing Pharmacy Association (ISPPA)	India
Raghu Ram	Nair	Symplifye	UK
Yang	Nan	Imperial College London	UK
Andrew	Narracott	University of Sheffield	UK
Marty	Nash	University of Auckland	New Zealand
Sandeepak	Natu	Trinity Technologies	Ireland
Majid	Nazemi	University of Liege	Belgium
Aurel	Neic	NumeriCor GmbH	Austria
Michael	Neidlin	RWTH Aachen University	Germany
Kelsy	Nguyen	IESEG school of management	France
Anna	Niarakis	University of Toulouse III-Paul Sabatier	France
David	Nickerson	University of Auckland	New Zealand
Steven	Niederer	King's College London	UK
Chris	Niklas	Heidelberg University	Germany
Christos	Nikolaou	Imperial College London	UK
Jerome	Noailly	Universitat Pompeu Fabra	Spain
Vincent	Noël	Institut Curie	France
Vit	Novacek	University of West Bohemia	Czech Republic
Jessica	Oakes	Northeastern University	USA
Susan	Ohanlon	Fiona Stanley Hospital	Australia
Rafael	Ojeda	Junta De Andalucia	Spain
Ivan	Olier-Caparroso	Liverpool John Moores University	UK
Simona	Olmi	Consiglio Nazionale delle Ricerche (CNR)	Italy

John	O'Shea	GeneEra	Ireland
Paivi	Ostling	SciLifeLab	Sweden
Marco	Ottella	Xtremion Technology GmbH	Germany
Frank	Otto	University of Amsterdam	The Netherlands
Özgür	Özdemir	Digital Factory	Turkey
Deniz	Ozturk	Fankom Engineering	Turkey
Susanna	Pakkasmaa	University of Bergen	Norway
Renata	Palen	European Commission	Luxembourg
Nuno	Palma dos Reis	NOVA Medical School	Portugal
Alexia	Papadimitriou	DIGITALEUROPE	Belgium
Francesco	Pappalardo	University of Catania	Italy
Georganta	Paraskevi	CMG LEGAL	France
Andrew	Parton	Multiomic Health	UK
Salvatore	Pasta	University of Palermo	Italy
Alex	Patak	European Commission, Joint Research Centre Directorate F –	Belgium
Gunther	Paul	James Cook University	Australia
Marie Helene	Pautrat	Inria	France
Morten	Pedersen	Università di Padova	Italy
Mathias	Peirlinck	TU Delft	The Netherlands
Esteban	Peña-Pitarch	Universitat Politècnica de Catalunya	Spain
María Ángeles	Pérez Ansón	University of Zaragoza	Spain
Florian	Petcu	Clinical Pediatric Hospital of Sibiu	Romania
Christophe	Phillips	University of Liège	Belgium
Gemma	Piella	Universitat Pompeu Fabra	Spain
Gabriele	Piergiovanni	Alient venture studio	Italy
Jens	Piesk	Nuromedia	Germany
Tommaso	Pippucci	IRCCS Azienda Ospedaliero-Universitaria di Bologna	Italy
Selene	Pirola	TU Delft	The Netherlands
Senol	Piskin	Istinye University, Istanbul	Turkey
Kelly	Plueschke	European Medicines Agency (EMA)	The Netherlands
Antonis	Polemitis	University of Nicosia	Cyprus
Maharajah	Ponnaiah	IHU ICAN	France
Aleksander	Popel	Johns Hopkins University	USA
Fernando	Portilla	Agder University	Norway
Mark	Potse	Inria	France
Saleh	Pouresmaeli	University of Sheffield	UK
Norman	Powell	University of Sheffield	UK
Drew	Pruett	HC Simulation	USA
Antonio	Puertas Gallardo	European Commission	Italy
Michiel	Punt	eScience Center	The Netherlands
Nicholas	Purcell	PurcellLtd	UK
Eleanor	Quek	Imperial College London	UK
Frank	Rademakers	KU Leuven	Belgium
Helge	Rader	University of Bergen	Norway
Laura	Rahm	LMU Munich	Germany
Ines	Ramos	Deloitte	Belgium

Giovanna	Ramos Queda	Université d'Aix Marseille	France
	Rampadarath	University of Auckland	New Zealand
Rajiv	Ranjan	DE INCUBATOR PTE. LTD.	Singapore
Andrea	Rappagliosi	Edwards	USA
Simon	Rapple	Karolinska Institutet	Sweden
John	Rasmussen	Aalborg University	Denmark
Magnificus	Rector	VU Amsterdam	The Netherlands
Francesco	Regazzoni	University of Amsterdam	The Netherlands
Markus	Rehberg	Sanofi	Germany
Kristin	Reiche	Fraunhofer Institute for Cell Therapy and Immunology	Germany
Alireza	Rezvani Sharif	CSL	Australia
Kinza	Rian	Junta de Andalucia	Spain
Wannes	Ribbens	KU Leuven	Belgium
Ron	Ribitzky	R&D Ribitzky	USA
John	Rice	Retired	USA
Mathieu	Rimaud	TwInsight	France
Petra	Ritter	Charité - Universitätsmedizin Berlin	Germany
Anna	Rizzo	Lynkeus	Italy
Lucia	Robson	Human cell atlas	UK
Maria	Rocchi	KU Leuven	Belgium
Helen	Roche	University College Dublin	Ireland
Michel	Rochette	Ansys	France
Cristobal	Rodero	King's College London	UK
Blanca	Rodriguez	University of Oxford	UK
Agustin	Rodriguez Gonzalez	German Cancer Research Center	Germany
Maria	Rodriguez Martinez	IBM Research	Switzerland
Julien	Rolland	Dassault Systèmes	France
Jolien	Roovers	Department Economy, Science & Innovation of the Flemish government	Belgium
Fariba	Roshanzamir	Chalmers university of technology	Sweden
Laurence	Rouet	Philips Ultrasound Research	France
Andreas	Rowald	University of Erlangen-Nuremberg	Germany
Sitikantha	Roy	IIT Delhi	India
Maria	Rujano	ECRIN	France
Gianpaolo	Ruocco	University of Basilicata and iBMB Srl	Italy
Ine	Rusten	Systems Resource Lab AS (SRLAB)	Norway
Christina	Saak	Universität Hamburg	Germany
Babak	Saboury	National Institutes of Health (NIH)	USA
Michael	Sacks	University of Texas at Austin	USA
Soroush	Safaei	University of Auckland	New Zealand
Oceane	Saibou	Université d'Évry	France
	Saiz	Universitat Politècnica de València	Spain
Md Nazmus	Sakib	University of Texas at Arlington	USA
Albin	Salazar	Ecole Normale Supérieure	France
Jelle	Saldien	IMEC – University of Ghent	Belgium
Raoul	Salle De Chou	Inria	France
Roser	Sanchez Todo	neuroelectronics	Spain

Venkata	Satagopam	University of Luxembourg	Luxembourg
Simone	Scacchi	University of Milan	Italy
Ella	Scallan	Leeds	UK
Stephan	Schaller	esqlabs	Germany
Ute	Schepers	Karlsruhe Institute of Technology, Karlsruhe	Germany
Silvia	Schievano	UCL & Great Ormond Street Hospital	UK
Winfried	Schlee	Eastern Switzerland University of Applied Sciences	Switzerland
Andreas	Schoenau	Agentur für Innovation in der Cybersicherheit GmbH	Germany
Patrick	Segers	Ghent University	Belgium
Timothy	Sego	University of Florida	USA
Martin	Sending	Oslo universitetssykehus HF	Norway
Marco	Sensale	Ansys	France
Benjamin	Senst	Freelance Data Scientist	Germany
Julia	Sergienko	InSysBio	UK
Marco	Seri	University of Bologna	Italy
Rahuman	Sheriff	European Bioinformatics Institute	UK
Amanda	Shick	University of Florida	USA
Isaac	Shiri	Bern University Hospital	Switzerland
Viktoriia	Shportiuk	Siemens Healthineers	Belgium
Sacha	Silva Saffar	University of Paris-Saclay	France
Jorge	Simoes	University of Twente	The Netherlands
Sourav	Sinha	EY Consulting	Belgium
Julianne	Skille	University of Bergen	Norway
James	Sluka	Indiana University	USA
Sylvain	Soliman	Inria	France
Karen	Sonego	City University of London	UK
Luis	Sordo Vieira	University of Florida	USA
Paul	Sovelius	McGovern Medical School	USA
Nicoleta	Spinu	Liverpool John Moores University	UK
Holger	Sprengel	Nuromedia	Germany
Bart	Spronck	Maastricht university	The Netherlands
Anil	Srivastava	Open Health Systems Laboratory	USA
Jim	St Clair	MyLigo, Inc	USA
Georgios	Stamatakis	National Technical University of Athens	Greece
Ernst-Jan	Stokvis	Stokvis consulting	The Netherlands
Lidia	Strigari	Ospedale Sant'Orsola Bologna	Italy
Jacob	Sturdy	Norwegian University of Science and Technology	Norway
Martina	Summer-Kutmon	Maastricht University	The Netherlands
Paul	Summers	QMRI Tech	Italy
Surash	Surash	TüvSüd	Germany
Audrius	Sveikata	BIOMAPAs	Lithuania
	Szepietowska	Gdańsk University of Technology	Poland
Marcela	Szopos	Paris Descartes University	France
Hugo	Talbot	Inria	France
Daniele	Tartarini	University of Sheffield	UK

Merryn	Tawhai	University of Auckland	New Zealand
Jack	Taylor	University of Sheffield	UK
Jesper	Tegner	KAUST & Karolinska Institutet	Singapore / Sweden
Sarah	Teichmann	Human cell atlas	UK
Jesper	Thestrup	In Jet	Denmark
Eric	Thibaut	IHU ICAN	France
Scott	Thiel	Hologic	USA
Paolo	Tieri	Consiglio Nazionale delle Ricerche (CNR)	Italy
Simone	Tischler	Julich	Germany
Alessandro	Tognetti	University of Pisa	Italy
Andreas	Tolk	MITRE	USA
Rudi	Torfs	VITO	Belgium
Manuel	Torres Rodrigues	SURF	The Netherlands
Alberto	Tozzi	Ospedale Pediatrico Bambino Gesù	Italy
Aline	Treuren	Erasmus MC	The Netherlands
Konstantinos	Triantos	Athena Research Center	Greece
Marco	Trombetti	PI School	Italy
Steven	Truijen	U Antwerpen	Belgium
Panagiotis	Tsakanikas	National Technical University of Athens	Greece
Eirini	Tsirvouli	Norwegian University of Science and Technology	Norway
Hervé	Turlier	Collège de France, CNRS	France
Mark	Turner	conect4children Stichting	The Netherlands
Francesco	Ubertini	CINECA	Italy
Stefan	Uhlig	Uniklinik RWTH Aachen	Germany
Mateo	Valero	Barcelona Supercomputing Center	Spain
Victoria	Valls	AQuAS -- Agency for Health Quality and Assessment of Catalonia	Spain
Aad	van de Lugt	Erasmus MC	The Netherlands
Frans	van de Vosse	Eindhoven University of Technology	The Netherlands
Robin	van der Aa	Erasmus MC	The Netherlands
Herman	Van der Auweraer	KU Leuven	Belgium
Johanna	van der Bom	Leiden University Medical Center	The Netherlands
Manon	van der Haeghe	KU Leuven	Belgium
Marcel	van der Kuil	Erasmus Medical Centre	The Netherlands
Andries	Van der Meer	University of Twente	The Netherlands
Geert	Van Gassen	Takeda Belgium	Belgium
Mark	Van Gils	University of Tampere	Finland
Ine	van Hoyweghen	KU Leuven	Belgium
Harry	van Lenthe	KU Leuven	Belgium
Rob	van Nieuwpoort	eScience Center	The Netherlands
Rick	van nuland	Lygature	The Netherlands
Frank	van Praat	KPMG-NL/AMdEX	The Netherlands
Natal	van Riel	Eindhoven University of Technology	The Netherlands
Lore	Van Santvliet	KU Leuven	Belgium
Kristel	Van Steen	University of Liège	Belgium
Jos	Vander Sloten	KU Leuven	Belgium
Laura	Vangeel	KU Leuven	Belgium

Pavlos	Varsos	Inria	France
Vasileios	Vavourakis	University of Cyprus	Cyprus
Jithesh	Veetil	Medical Device Innovation Consortium	USA
Alan	Velazquez	École Supérieure d'Ingénieurs Léonard de Vinci	France
Alexey	Velikorodnyy	BIOME Science	France
Peter-Paul	Verbeek	Universiteit van Amsterdam	The Netherlands
Marco	Verdicchio	SURF	The Netherlands
Nico	Verdonschot	Radboud University Medical Center	The Netherlands
Natalie	Vermeulen	European Society of Human Reproduction and Embryology	Belgium
Lorenzo	Veschini	King's College London	UK
Eric	Vibert	Greater Hospital in Paris AP-HP	France
Gabriele	Vicedomini	Policlinico San Donato - San Donato Milanese	Italy
Edward	Vigmond	Université Bordeaux	France
Marco	Virgolin	InSilicoTrials Technologies SPA	Italy
Bruno	Virieux	Predisurge	France
Tapio	Visakorpi	University of Tampere	Finland
Luigi	Visani	AIMED trial srl	Italy
Arun	Viswanathan	Dassault Systèmes	USA
Ana	Vlasceanu	Inria	France
Jahn	Volock	Medtronic	The Netherlands
Falk	Von Dincklage	Universitätsmedizin Greifswald	Germany
Veronika	Vsetickova	European Liver Patients' Association	Belgium
Britta	Wagenhuber	Sanofi	Germany
William	Waites	University of Strathclyde, Glasgow	UK
Melody	Walker	University of Florida	USA
Dagmar	Waltemath	University of Greifswald	Germany
Berend	Westerhof	Westerhof Cardiovascular Research	The Netherlands
Maurice	Whelan	European Commission, Joint Research Center, JRC.F	Italy
Carl	Whitfield	University of Manchester	UK
Tobias	Wiesenthal	European Commission, Joint Research Center, JRC.F	Belgium
James	Wild	University of Sheffield	UK
Matthew	Williams	Robotech	India
Nathaniel	Wilson	International School	Hong Kong (SAR)
Willi	Wolfarth	University Heidelberg	Germany
Jelmer	Wolterink	University of Twente	The Netherlands
Yinzhe	Wu	Imperial College London	UK
Vasileios	Xanthakis	Karlsruhe Institute of Technology	Germany
Guang	Yang	Imperial College London	UK
Justin	Yea	Medtronic Korea Ltd.	Korea
Funda	Yildirim	University of Twente	The Netherlands
Gabriel	Zachmann	University of Bremen	Germany
Elisabetta	Zanette	CINECA	Italy
Veronica	Zarnitsyna	Emory University Atlanta GA	USA

Laurynas	Zemaitaitis	Innovative Medicine Centre	Lithuania
Naouel	Zerrouk	Université d'Évry	France
Begum	Zeybek	Teesside University	UK
Debbie	Zhao	University of Auckland	New Zealand
Vadim	Zhernovkov	University College Dublin	Ireland
Hong	Zhu	Soochow University	China
Andrei	Zinovyev	Institut Curie	France
Antonio	Zoccoli	INFN Istituto Nazionale Fisica Nucleare	Italy
Silvia	Zullo	University of Bologna	Italy
Paolo	Zunino	Politecnico di Milano	Italy

37 Annex 2: Standards and standard defining organisations

The EDITH FAIRsharing standards collection⁵⁶³ contains an interactively browsable and searchable online listing of the standards mentioned in the tables below and will be kept up to date.

Table A 12: Technical standard defining organisations

Standard defining organisation	Description
ASME ⁵⁶⁴	American Society of Mechanical Engineers Defines codes and standards for mechanical engineering. Important subcommittees are: - VVUQ SC 40 Computational Modelling of Medical Devices - VVUQ SC 60 Guideline for Simulation Software Selection - VVUQ SC 70 Machine Learning Applied to Mechanistic & Process Modelling
CEN / CENELEC ⁵⁶⁵	European Committee for Standardisation The CEN Technical Committee CEN/TC 251 defines standards for health informatics
COMBINE ⁵⁶⁶	Computational modelling in Biology Network. Coordinates the development of community standard formats for systems biology modelling.
DICOM ⁵⁶⁷	Digital Medicine and Communications in Medicine DICOM is a standard for medical imaging data (ISO 12052:2017).
GA4GH ⁵⁶⁸	Genome Alliance for Genomics and Health Standards for collecting, storing, analysing, and sharing of genomic data within research and healthcare
HL7 ⁵⁶⁹	Health Level 7 (since it is located on the highest level 7, the application layer of the ISO/OSI model communication model) Defined a set of standards for the exchange of electronic clinical and health administrative information
IEC ⁵⁷⁰	International Electrotechnical Commission
IEEE ⁵⁷¹	Institute of Electrical and Electronics Engineers
ISO/TC 215 ⁵⁷²	International Standards organisation (ISO) Technical Committee 215 Standardisation in the field of health and medical informatics
ISO/TC 276/WG5 ⁵⁷³	International Standards organisation (ISO) Technical Committee 276 (meta)data standards for the life sciences and for data processing
ICHOM ⁵⁷⁴	Standards organisation for patient outcomes

⁵⁶³ <https://fairsharing.org/4787>

⁵⁶⁴ <https://www.asme.org/codes-standards>

⁵⁶⁵ <https://standards.cenelec.eu>

⁵⁶⁶ <http://co.mbine.org>

⁵⁶⁷ <https://www.dicomstandard.org>

⁵⁶⁸ <https://www.ga4gh.org>

⁵⁶⁹ <https://www.hl7.org>

⁵⁷⁰ <https://www.iec.ch/homepage>

⁵⁷¹ <https://standards.ieee.org>

⁵⁷² <https://www.iso.org/committee/54960.html>

⁵⁷³ <https://www.iso.org/committee/4514241.html>

⁵⁷⁴ <https://www.ichom.org/aboutus/>

Table A 13: Clinical standard defining organisations

Standard defining organisation	Description
CDISC ⁵⁷⁵	Clinical Data Interchange Standards Consortium Definition of several standards for clinical trials and case report forms (<i>e.g.</i> , ADaM, CTR-XML, ODM-XML, OMOP, SDTM)
C-Path ⁵⁷⁶	Critical Path Institute
ECRI	Emergency Care Research Institute
ICH ⁵⁷⁷	International Council for Harmonisation
IDMP ⁵⁷⁸	Identification of Medicinal Products
IHE	Integrating the Healthcare Enterprise
IHTSDO ⁵⁷⁹	International Health Terminology Standards Development Organisation, the organisation defining the SNOMED-CT ⁵⁸⁰ (Systematised Nomenclature of Medicine – Clinical Terms) terms
NCI-EVS ⁵⁸¹	National Cancer Institute - Enterprise Vocabulary Service
OHDSI ⁵⁸²	Observational Health Data Sciences and Informatics
Regenstrief ⁵⁸³	The Regenstrief institute defines the LOINC and UCUM codes
TransCelerate ⁵⁸⁴	TransCelerate Clinical Data Standards Initiative
WHO	World Health Organisation

⁵⁷⁵ <https://www.cdisc.org/standards>⁵⁷⁶ <https://c-path.org/news-and-events/data-standards/>⁵⁷⁷ <https://www.ich.org/page/ich-guidelines>⁵⁷⁸ <https://www.fda.gov/industry/fda-data-standards-advisory-board/identification-medicinal-products-idmp>⁵⁷⁹ <https://confluence.ihtsdotools.org>⁵⁸⁰ <https://www.snomed.org>⁵⁸¹ <https://evs.nci.nih.gov>⁵⁸² <https://www.ohdsi.org>⁵⁸³ <https://www.regenstrief.org>⁵⁸⁴ <https://www.transceleratebiopharmainc.com/initiatives/clinical-data-standards/>

Table A 14: Standards for medical imaging

Medical standard	imaging	Description																		
BIDS ⁵⁸⁵		Brain Imaging Data Structure, a standard of the INCF ⁵⁸⁶ (International Neuroinformatics Coordinating Facility) for capturing data and metadata of MRI datasets. For different imaging techniques the following extensions ⁵⁸⁷ to BIDS are available: <table><tr><td>ASL-BIDS</td><td>- for arterial spin labelling</td></tr><tr><td>EEG-BIDS</td><td>- for electroencephalographic data</td></tr><tr><td>iEEG-BIDS</td><td>- for intracranial EEG data</td></tr><tr><td>fNIRS-BIDS</td><td>- for Near InfraRed Spectroscopy</td></tr><tr><td>Genetics-BIDS</td><td>- for genetic data associated with human brain imaging</td></tr><tr><td>MEG-BIDS</td><td>- for magnetoencephalographic data</td></tr><tr><td>Microscopy-BIDS</td><td>- for microscopy imaging data</td></tr><tr><td>PET-BIDS</td><td>- for positron emission tomography</td></tr><tr><td>qMRI-BIDS</td><td>- for quantitative MRI data</td></tr></table>	ASL-BIDS	- for arterial spin labelling	EEG-BIDS	- for electroencephalographic data	iEEG-BIDS	- for intracranial EEG data	fNIRS-BIDS	- for Near InfraRed Spectroscopy	Genetics-BIDS	- for genetic data associated with human brain imaging	MEG-BIDS	- for magnetoencephalographic data	Microscopy-BIDS	- for microscopy imaging data	PET-BIDS	- for positron emission tomography	qMRI-BIDS	- for quantitative MRI data
ASL-BIDS	- for arterial spin labelling																			
EEG-BIDS	- for electroencephalographic data																			
iEEG-BIDS	- for intracranial EEG data																			
fNIRS-BIDS	- for Near InfraRed Spectroscopy																			
Genetics-BIDS	- for genetic data associated with human brain imaging																			
MEG-BIDS	- for magnetoencephalographic data																			
Microscopy-BIDS	- for microscopy imaging data																			
PET-BIDS	- for positron emission tomography																			
qMRI-BIDS	- for quantitative MRI data																			
DICOM ⁵⁸⁸		Digital Imaging and Communications in Medicine, a standard for medical imaging-based modelling; stores the data as 2D layers																		
CIFTI-2 ⁵⁸⁹		Connectivity Informatics Technology Initiative (grayordinate surface + volume)																		
GIFTI ⁵⁹⁰		Geometry (surface) file format																		
NIFTI-2 ⁵⁹¹		Neuroimaging Informatics Technology Initiative: 2nd Version of a special image format for neuroimaging, where the data are stored in a true 3D volume format																		
NIFTI-MRS ⁵⁹²		extension of NIFTI for magnetic resonance spectroscopy																		

⁵⁸⁵ <https://bids.neuroimaging.io>⁵⁸⁶ <https://www.incf.org>⁵⁸⁷ https://bids.neuroimaging.io/get_involved.html#extending-the-bids-specification⁵⁸⁸ <https://www.dicomstandard.org>⁵⁸⁹ https://www.nitrc.org/forum/attachment.php?attachid=333&group_id=454&forum_id=1955⁵⁹⁰ <https://www.nitrc.org/projects/gifti/>⁵⁹¹ <https://nifti.nih.gov/nifti-2>⁵⁹² https://wtclarke.github.io/mrs_nifti_standard/

Table A 15: Standard formats for electro- and neurophysiology, biosignal and vital sign data

Standard	Description
ECG-XML	An XML-based format for electrocardiogram data
EDF / EDF+ ⁵⁹³	European Data Format, a format for exchange and storage of multichannel biological and physical signals, <i>e.g.</i> , polysomnography
GDF ⁵⁹⁴	General Data Format; for biomedical signals, like <i>e.g.</i> , EEG and ECG data
HL7-aECG ⁵⁹⁵	HL7-annotated Electrocardiogram; an annotated XML-based format for electrocardiogram data
ISO/IEEE 11073 ⁵⁹⁶	A standard for device interoperability of point-of-care medical device communication
MEF ⁵⁹⁷	Multiscale Electrophysiology Format for EEG data
MFER ⁵⁹⁸	Medical waveform Format Encoding Rules, a file format for encoding medical waveforms from ECG, ECoG and EEG data
NIX ⁵⁹⁹	Neuroscience Information eXchange for neurophysiology data
NSDF ⁶⁰⁰	Neuroscience Simulation Data Format based on HDF5
NWB:N 2.0 ⁶⁰¹	Neurodata Without Borders:Neurophysiology for neurophysiology data
OpenEP ⁶⁰²	A cross-platform electroanatomic mapping data format
SCP-ECG ⁶⁰³	Standard Communications Protocol for ECG data
SignalML ⁶⁰⁴	An XML-based meta-format for biomedical time series data
SONATA ⁶⁰⁵	Scalable Open Network Architecture TemplAte; for large-scale modelling of neuronal brain network models and simulation output
VSIR ⁶⁰⁶	Vital Signs Information Representation (CEN 13734)
WFDB ⁶⁰⁷	WaveForm DataBase format of the PhysioNet (physio.net) repository; a combination of MIT format and EDF / EDF+

⁵⁹³ <https://www.edfplus.info>⁵⁹⁴ http://justsolve.archiveteam.org/wiki/General_Data_Format_for_Biosignals⁵⁹⁵ https://en.wikipedia.org/wiki/HL7_aECG⁵⁹⁶ <https://www.iso.org/standard/77338.html>⁵⁹⁷ <https://main.ieeg.org/?q=node/28>⁵⁹⁸ <http://www.mfer.org/en/index.htm>⁵⁹⁹ <http://g-node.github.io/nix/>⁶⁰⁰ <https://github.com/nsdf/nsdf>⁶⁰¹ <https://www.nwb.org/2019/02/26/nwbn-2-0-final-released/>⁶⁰² <https://openep.io>⁶⁰³ <https://github.com/topics/scp-ecg>⁶⁰⁴ <https://braintech.pl/software/svarog/signalml/?lang=en>⁶⁰⁵ <https://docs.sonata-project.org/en/master/>⁶⁰⁶ <https://standards.iteh.ai/catalog/standards/cen/8f621d17-ffcc-4885-bfb2-ff72df02e7f1/env-13734-2000>⁶⁰⁷ <https://wfdb.readthedocs.io/en/latest/wfdb.html>

Table A 16: Standards for genetic sequence variants

Sequence format	variant	Description
VCF ⁶⁰⁸		Variant Call Format
BCF		Binary Call Format, a binary version of VCF
MAF ⁶⁰⁹		Mutation Annotation Format
GVF		Genome Variation Format, an extension of GFF3
GVCF ⁶¹⁰		Genomic Variant Call Format
SPDI		Sequence, Position, Deletion, Insertion
GA4GH-VR ⁶¹¹		Genome Alliance for Genomics and Health - Variation Representation
HGVS ⁶¹²		Human Genome Variation Society, following the HGVS sequence variant nomenclature

⁶⁰⁸ <https://gatk.broadinstitute.org/hc/en-us/articles/360035531692-VCF-Variant-Call-Format>

⁶⁰⁹ https://docs.gdc.cancer.gov/Data/File_Formats/MAF_Format/

⁶¹⁰ <https://gatk.broadinstitute.org/hc/en-us/articles/360035531812-GVCF-Genomic-Variant-Call-Format>

⁶¹¹ <https://vrs.ga4gh.org/en/latest/>

⁶¹² <https://varnomen.hgvs.org/bg-material/simple/>

Table A 17: Standards for models

Modelling standard	Description
BioPAX ⁶¹³	Biological Pathways eXchange for exchange and visualisation of biological pathway data
CellML ⁶¹⁴	XML-based description and exchange format for cellular models
FieldML ⁶¹⁵	Human Physiome Field Markup Language, an XML-based format using mathematical field descriptions of cells, tissues and organs; can be used to represent finite element models
MDL ⁶¹⁶	Model Description Language for pharmacometric models
MoBi ⁶¹⁷	For multiscale physiological modelling and simulation
MorpheusML ⁶¹⁸	Format for agent-based multicellular models
MultiCellIDS ⁶¹⁹	MultiCellular Data Standard for centre-based models (CBMs) [Montagud 2021]
MultiCellML ⁶²⁰	Standard for agent-based multiscale and multicellular spatial models
NeuroML v2 ⁶²¹	Neuroscience eXtensible Markup Language; an XML-based description and exchange format for models in neuroscience with its four parts: Biophysics, ChannelML, MorphML, and NetworkML
OpenBEL ⁶²²	Biological Expression Language, a triple-based (subject-predicate-object) modelling language for representing biological knowledge by causal, correlative, and associative relationships
PharmML	Pharmacometrics Markup Language, an exchange format for pharmacokinetic and pharmacodynamic models
PK-Sim ⁶²³	Standard model for whole-body physiologically based pharmacokinetic modelling
SBML ⁶²⁴	Systems Biology Markup Language, an XML-based description and exchange format for differential-equation models of biological processes. SBML Level 3 is a modular format with a core and packages for extending that core functionality:
	SBML-arrays - for vectorized, <i>e.g.</i> , grid-based models
	SBML-comb - for multiscale and modular, <i>e.g.</i> , tissue and whole-body models
	SBML-distrib - for distributions, <i>e.g.</i> , systems pharmacology and population models
	SBML-dyn - for dynamical models
	SBML-fbc - for constraint-based flux-balance (steady state), <i>e.g.</i> , genome-scale models
	SBML-multi - for rule-based, <i>e.g.</i> , multistate molecules and multicomponent complexed
	SBML-layout - for visualisation
	SBML-groups - for grouping and organisation
	SBML-qual - for qualitative (<i>i.e.</i> , Boolean) models
	SBML-render - for visualisation
	SBML-spatial - for spatial, <i>e.g.</i> , reaction-diffusion models

⁶¹³ <http://www.biopax.org>⁶¹⁴ <https://www.cellml.org>⁶¹⁵ <https://physiomeproject.org/software/fieldml>⁶¹⁶ <http://mdl.community>⁶¹⁷ <https://github.com/Open-Systems-Pharmacology/MoBi>⁶¹⁸ <https://morpheus.gitlab.io/tag/morpheusml/>⁶¹⁹ <http://multicellids.org>⁶²⁰ <https://multicellml.org/wiki/doku.php>⁶²¹ <https://neuroml.org>⁶²² <https://github.com/OpenBEL>⁶²³ <https://github.com/Open-Systems-Pharmacology/PK-Sim>⁶²⁴ <https://sbml.org>

SBOL ⁶²⁵	Synthetic Biology Open Language, describing the exchange of synthetic biological genetic parts, devices, modules, and systems.
VCML ⁶²⁶	Virtual Cell Markup Language for rule-based modelling

⁶²⁵ <https://sbolstandard.org>

⁶²⁶ https://vcell.org/webstart/VCell_Tutorials/VCell_Help/topics/ch_1/Introduction/Export.html

Table A 18: Standards for model simulations and documentation of results

Standard	Description
GSP ⁶²⁷	Good Simulation Practice, a quality standard for <i>in silico</i> simulations, developed by the Avicenna Alliance
NuML ⁶²⁸	Numerical Markup Language, an XML-based format for exchanging numerical data
OMEX	Open modelling EXchange for exchange of modelling and simulation data. An OMEX file is a .zip container containing a manifest file, an optional metadata file and the data files.
SBRML ⁶²⁹	Systems Biology Results Markup Language for encoding results of SBML simulations
SED-ML ⁶³⁰	Simulation Experiment Description, a XML-based exchange format for encoding of simulation setups following the MIASE guidelines

⁶²⁷ <https://insilico.world/sito/wp-content/uploads/2023/05/Position-Paper-GSP-R6.pdf>

⁶²⁸ <https://github.com/NuML/NuML>

⁶²⁹ <https://sbrml.sourceforge.net/SBRML/Welcome.html>

⁶³⁰ <https://sed-ml.org>

Table A 19: Terminologies and ontologies for the description and annotation of data, models, and their components

Ontology	Description
BCTEO	Bone/Cartilage Tissue Engineering Ontology
BRICKS ⁶³¹	Describes recurring concept in SBGN models
CBO ⁶³²	The Cell Behavior Ontology is designed to describe multi-cell computational models
CheBI ⁶³³	Chemical Entities of Biological Interest
CMDO	Clinical MetaData Ontology
CNO ⁶³⁴	Computational Neuroscience Ontology
CO	Cell Ontology
COMODI ⁶³⁵	COMputational MOdels Differ, an ontology to characterise differences in versions of computational biology Models
CVDO	Cardiovascular Disease Ontology
DEB ⁶³⁶	Devices, Experimental Scaffolds, and Biomaterials Ontology
DINTO ⁶³⁷	Drug-drug Interaction Ontology
DO ⁶³⁸	Disease ontology
DTO ⁶³⁹	Drug target ontology
DUO ⁶⁴⁰	Data Use Ontology
ECTO ⁶⁴¹	Environmental Conditions, Treatments, and exposures Ontology
EDAM ⁶⁴²	EMBRACE Data and Methods ontology
EnzymeML ⁶⁴³	XML-based format for storage and transfer of enzyme kinetics data
FMA ⁶⁴⁴	Foundational Model of Anatomy for localising anatomical structures at a specific spatial location
HPO ⁶⁴⁵	Human Phenotype Ontology
HUPSON	Human Physiology Simulation Ontology containing concepts for simulations, models, and algorithms
InChI ⁶⁴⁶	IUPAC International Chemical Identifier
KISAO ⁶⁴⁷	Kinetic Simulation Algorithm Ontology for describing simulation algorithms
LiCO	Liver case ontology
MAMO	Mathematical Modelling Ontology
MAxO ⁶⁴⁸	Medical Action Ontology
MONDO ⁶⁴⁹	A semi-automatically constructed ontology defined by the Monarch initiative that merges multiple disease ontologies to yield a coherent merged ontology.
ND ⁶⁵⁰	Neural Disease Ontology

⁶³¹ <https://brickschema.org/ontology/>

⁶³² <https://cbo.biocomplexity.indiana.edu>

⁶³³ <https://www.ebi.ac.uk/chebi>

⁶³⁴ <https://github.com/INCF/Computational-Neurosciences-Ontology--C.N.O.->

⁶³⁵ <http://comodi.sems.uni-rostock.de>

⁶³⁶ https://github.com/ProjectDebbie/Ontology_DEB

⁶³⁷ <https://github.com/labda/DINTO>

⁶³⁸ <https://disease-ontology.org>

⁶³⁹ <http://drugtargetontology.org>

⁶⁴⁰ <https://github.com/EBISPOT/DUO>

⁶⁴¹ <http://obofoundry.org/ontology/ecto.html>

⁶⁴² <https://edamontology.org/page>

⁶⁴³ <https://enzymeml.org>

⁶⁴⁴ <http://sig.biostr.washington.edu/projects/fm/AboutFM.html>

⁶⁴⁵ <https://hpo.jax.org/app/>

⁶⁴⁶ <https://iupac.org/who-we-are/divisions/division-details/inchi/>

⁶⁴⁷ <https://github.com/SED-ML/KiSAO/>

⁶⁴⁸ <https://github.com/monarch-initiative/MAxO>

⁶⁴⁹ <https://mondo.monarchinitiative.org>

⁶⁵⁰ <https://github.com/addiehl/neurological-disease-ontology>

NPO ⁶⁵¹	NanoParticle Ontology
NPU	Nomenclature for Properties and Units terminology
OBCS ⁶⁵²	Ontology of Biological and Clinical Statistics
OBI ⁶⁵³	Ontology for Biomedical Investigations
OGMS ⁶⁵⁴	Ontology for General Medical Science
OPD	Ontology of Physics for Biology containing physics concepts to describe the dynamics of biological systems
OPMI ⁶⁵⁵	Ontology of Precision Medicine and Investigation
PAV ⁶⁵⁶	Ontology for provenance, authoring and versioning
PK ontology	An ontology for pharmacokinetic models
PreMedOnto	Precision Medicine Ontology
ROO ⁶⁵⁷	Radiation Oncology Ontology
SBO ⁶⁵⁸	Systems Biology Ontology with terms useful for describing computational modelling
SBOL-VO	Synthetic Biology Open Language Visual Ontology
SIO ⁶⁵⁹	Semantic science Integrated Ontology
STATO ⁶⁶⁰	Ontology of Statistical methods
Strenda ⁶⁶¹	Standard for reporting enzymology data
TEDDY	Terminology for the Description of Dynamics. An Ontology for the description of control elements and dynamics in systems biology and synthetic biology
UCUM ⁶⁶²	Unified Code for Units of Measure, defined by the Regenstrief institute

⁶⁵¹ <http://www.nano-ontology.org>

⁶⁵² <https://github.com/obcs/obcs>

⁶⁵³ <https://obi-ontology.org>

⁶⁵⁴ <https://github.com/OGMS>

⁶⁵⁵ <https://github.com/OPMI/opmi>

⁶⁵⁶ <https://pav-ontology.github.io/pav/>

⁶⁵⁷ <https://www.cancerdata.org/roo-information>

⁶⁵⁸ <https://github.com/EBI-BioModels/SBO>

⁶⁵⁹ <https://github.com/MaastrichtU-IDS/semanticscience>

⁶⁶⁰ <https://stato-ontology.org>

⁶⁶¹ <https://www.beilstein-institut.de/en/projects/strenda>

⁶⁶² <https://ucum.nlm.nih.gov>

Table A 20: Clinical languages, terminologies, and code systems

Language, terminology / code system	Description
Arden syntax ⁶⁶³	a HL7 standard for the representation of medical knowledge, <i>e.g.</i> , for use by CDSSs
ATC ⁶⁶⁴	Anatomical Therapeutic Chemical code, a classification system for the active substances of biomedical drugs
CQL ⁶⁶⁵	Clinical Quality Language, a HL 7 standard for the expression of clinical knowledge used for CDS and electronic Clinical Quality Measurement (eCQM); can also be used for querying complementing FHIR search
EMDN ⁶⁶⁶	European Medical Device Nomenclature
GMDN ⁶⁶⁷	Global Medical Device Nomenclature
ICD-11 ⁶⁶⁸	International Classification of Diseases
ICF ⁶⁶⁹	International Classification of Functioning, Disability and Health
ICHI ⁶⁷⁰	International Classification of Health Interventions
LOINC ⁶⁷¹	Logical Observation Identifiers Names and Codes for reporting laboratory test results
MedDRA ⁶⁷²	Medical Dictionary for Regulatory Activities
NCIt ⁶⁷³	the National Cancer Institute thesaurus
ORDO ⁶⁷⁴	Orphanet Rare Disease Ontology
ORPHAcode ⁶⁷⁵	Encoding of rare diseases and orphan drugs
RxNorm ⁶⁷⁶	Normalised Names for clinical drugs
SNOMED-CT ⁶⁷⁷	Systematized Nomenclature of Medicine – Clinical Terms
TNM ⁶⁷⁸	Tumour Node Metastasis, a classification system for malignant tumours
UMDNS ⁶⁷⁹	Universal Medical Device Nomenclature System of ECRI institute

⁶⁶³ https://www.hl7.org/implement/standards/product_brief.cfm?product_id=2

⁶⁶⁴ <https://www.ema.europa.eu/en/glossary/atc-code>

⁶⁶⁵ <https://cql.hl7.org>

⁶⁶⁶ <https://webgate.ec.europa.eu/dyna2/emdn/>

⁶⁶⁷ <https://www.gmdnagency.org>

⁶⁶⁸ <https://icd.org>

⁶⁶⁹ <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>

⁶⁷⁰ <https://www.who.int/standards/classifications/international-classification-of-health-interventions>

⁶⁷¹ <https://loinc.org>

⁶⁷² <https://www.meddra.org/>

⁶⁷³ <https://ncithesaurus.nci.nih.gov/ncitbrowser/>

⁶⁷⁴ <https://www.orpha.net/consor/cgi-bin/index.php?lng=EN>

⁶⁷⁵ <https://www.orpha.net/consor/cgi-bin/Disease.php?lng=EN>

⁶⁷⁶ <https://www.nlm.nih.gov/research/umls/rxnorm/overview.html>

⁶⁷⁷ <https://www.snomed.org>

⁶⁷⁸ <https://www.uicc.org/resources/tnm>

⁶⁷⁹ <https://www.ecri.org/solutions/umdns>

Table A 21: General metadata standards, formats, and protocols

CERIF ⁶⁸⁰	Common European Research Information Format <ul style="list-style-type: none"> • A conceptual model describing the Research domain. • defined as Entity Relationship Model (ERM).
DCAT ⁶⁸¹ -AP	Data CATalogue 3 – Application Profile: is a RDF-vocabulary defined by the W3C to facilitate interoperability between data catalogues on the web.
healthDCAT-AP ⁶⁸²	Metadata standard for supporting interoperability of health data in the EHDS ⁶⁸³ (European Health DataSpace). The HealthDCAT-AP editor ⁶⁸⁴ is a tool for creating datasets compliant to healthDCAT-AP.
METS ⁶⁸⁵	Metadata Encoding and Transmission Standard <ul style="list-style-type: none"> • for encoding descriptive, administrative, and structural metadata regarding objects within a digital library. • defined as XML schema.
OAI-PMH ⁶⁸⁶	Open Archives Initiative Protocol for Metadata Harvesting <ul style="list-style-type: none"> • A protocol developed for harvesting metadata descriptions of records in an archive to achieve repository interoperability. • defined as a set of six web requests or HTML verbs: GetRecord, Identify, ListIdentifiers, ListMetadataFormats, ListRecords, ListSets
XMP ⁶⁸⁷	Extensible Metadata Platform (ISO 16684-1:2019) <ul style="list-style-type: none"> • contains the following: <ul style="list-style-type: none"> - Data model: the structure of statements that XMP can make about resources. - Serialization: how any instance of the XMP data model can be recorded as XML • the embedding of XMP packets in specific file formats and domain specific XMP properties can be freely defined.
OpenAIRE Metadata format ⁶⁸⁸	Open Access Infrastructure for Research in Europe <ul style="list-style-type: none"> • expects the metadata to be encoded in the DataCite metadata format. • uses the OAI-PMH v2.0 protocol for harvesting dataset metadata.
DataCite Metadata Schema V4.4 ⁶⁸⁹	For publication and citation of research data <ul style="list-style-type: none"> • A list of core metadata properties chosen for an accurate and consistent identification of a resource for citation and retrieval purposes. • The metadata of Invenio357 are aligned with the DataCite metadata schema.

⁶⁸⁰ https://eurocris.org/eurocris_archive/cerifsupport.org

⁶⁸¹ <https://www.w3.org/TR/vocab-dcat-3>

⁶⁸² <https://healthdcat-ap.github.io>

⁶⁸³ <https://www.european-health-data-space.com>

⁶⁸⁴ <http://fair.healthdataportal.eu/editor2>

⁶⁸⁵ <https://www.loc.gov/standards/mets/mets-home.html>

⁶⁸⁶ <https://www.openarchives.org/pmh>

⁶⁸⁷ <https://www.iso.org/standard/75163.html>

⁶⁸⁸ <http://www.openaire.eu>

⁶⁸⁹ <http://schema.datacite.org>