Consortium Proposal

for the National Research Data Infrastructure (NFDI)

NFDI4Memory

Consortium for the historically oriented humanities

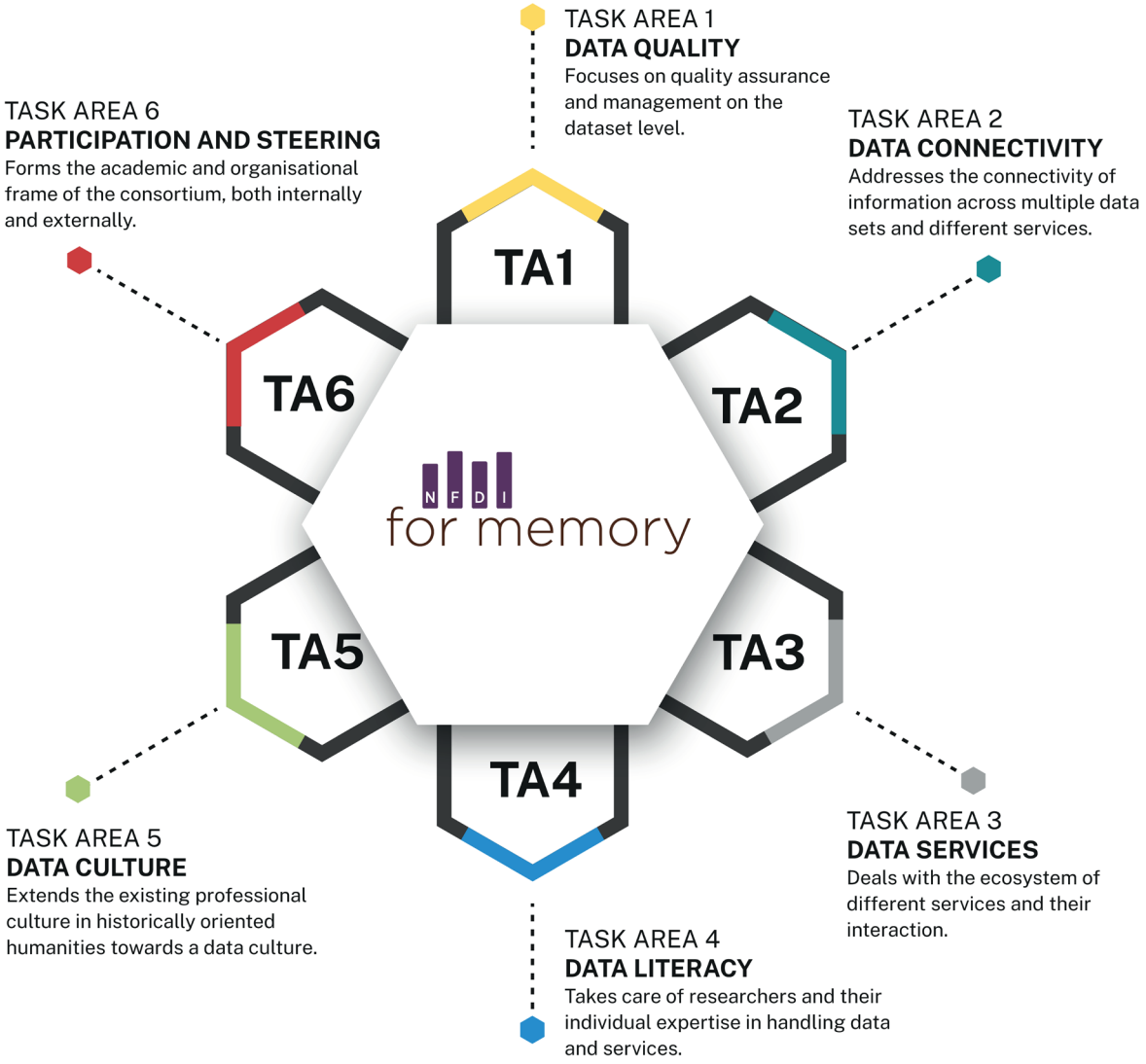

Table of Contents

Volume 1: Proposal

1 General Information 3
2 Scope and Objectives 14
2.1 Research domains or research methods addressed by the consortium, specific aim(s) .. 20
2.2 Objectives and measuring success 21
3 Consortium 23
3.1 Composition of the consortium and its embedding in the community of interest 23
3.2 The consortium within the NFDI 30
3.3 International networking 32
3.4 Organisational structure and viability 34
3.5 Operating model 36
4 Research Data Management Strategy 37
4.1 State of the art and needs analysis 37
4.2 Metadata standards 42
4.3 Implementation of the FAIR principles and data quality assurance 46
4.4 Services provided by the consortium 52
5 Work Programme 57
5.1 Task Area 1: Data Quality 58
5.2 Task Area 2: Data Connectivity 68
5.3 Task Area 3: Data Services 77
5.4 Task Area 4: Data Literacy 88
5.5 Task Area 5: Data Culture 97
5.6 Task Area 6: Participation and Steering 108

List of Abbreviations 122
Appendix 128
1 Bibliography and list of references 128

Citation

Paulmann, Johannes, John Wood, Klaus Ceynowa, Fabian Cremer, Silvia Daniel, Daniel Fähle, Barbara Fichtl, Peter Haslinger, Torsten Hiltmann, Rüdiger Hohls, Ursula Lehmkuhl, Marina Lemaire, Gerald Maier, Ole Meiners, Gisela Minn, Katrin Moeller, Andreas Neuburger, Claudia Prinz, Matthias Razum, Harald Sack, Johannes Sauter, Hildegard Schäffler, Eva Schlotheuber, Stefan Schmunk, Arnost Stanzel, Helmuth Trischler and Thorsten Wübbena: “NFDI4Memory. Consortium for the historically oriented humanities. Proposal for the National Research Data Infrastructure (NFDI)”, zenodo, 2022, DOI: https://doi.org/10.5281/zenodo.7428489.

List of Illustrations

Figure 1: 4Memory project design *cover*
Figure 2: 4Memory governance model *p. 34*
Figure 3: 4Memory service portfolio *p. 52*

1 General Information

Name of the consortium

NFDI4Memory – The Consortium for the Historically Oriented Humanities
NFDI4Memory – Konsortium für historisch arbeitende Geisteswissenschaften

Summary of the proposal in English

NFDI4Memory will set the pace for institutional, cultural and methodological innovation in historical research, for the first time **integrating research, memory and information infrastructure institutions** into a common RDM infrastructure and enabling a **sustainable transformation** of the historically oriented humanities. With our unique expertise in digital historical source criticism and hermeneutics, we drive forward the future knowledge order for the NFDI. Our innovative, sustainable RDM methods and services will enable a broad cultural change. The consortium unites established, pioneering co-applicants and participants from across the epochal breadth of historical research with strong experience in digital methods.

We have two connected aims. First, the growth of born-digital data and cutting-edge digital methods means that **entirely new areas of historical research are being opened**, requiring improved RDM infrastructures and services. 4Memory will advance historical research by providing RDM guidelines and digital services for a heterogenous research community still at an early stage of digital transformation. We will identify key innovations, evaluate their value to our community, foster their development, integrate them into existing structures and processes and promote their wider adoption. Our second aim is to **bridge a methodological gap** by bringing digital and analogue methodologies together. Digital transformation requires capacity building and cultural change, aims we pursue by enhancing digital competencies for those studying at universities, pursuing research, and working in archives, libraries and collections. 4Memory ensures the widest possible development of digital competencies and the creation of incentives and structures to make them a priority, thereby innovating historical scholarship.

We will create a **portfolio of key services** for historical research, building a 4Memory research data infrastructure that systematically links scholarship and digital resources. Our carefully integrated **internationalisation strategy** will ensure global RDM perspectives are brought into the defining discussions and procedures; unite international partners in research, memory and information infrastructure institutions; and strengthen essential cooperation with transnational structures and governing bodies. We will address and provide guidance on **legal and ethical issues** relevant to historians, and by developing scholarship to meet the challenges and opportunities of digitalisation, we will **safeguard the critical functions and social relevance** of the historical method. 4Memory is the only consortium focused on historical data and methods and is well placed to reflect upon the history of knowledge structures and, therefore, to place **digitalisation in a critical, long-term perspective** that will serve the NFDI as a whole.

Summary of the proposal in German

NFDI4Memory ist Schrittmacher institutioneller, fachlicher und methodischer Innovationen: Das Konsortium **integriert erstmals historische Forschungs-, Gedächtnis- und Informationsinfrastrukturen** in einer digitalen Forschungsdateninfrastruktur und ermöglicht durch die Entwicklung innovativer RDM-Methoden und Services die **nachhaltige Transformation** historisch arbeitender Geisteswissenschaften. Das Konsortium wird getragen von ausgewiesenen Co-Applicants mit großer Erfahrung in digitalen Methoden und von Participants aus der gesamten disziplinären und epochalen Breite der historischen Forschung. Ihre Expertise in digitaler historischer Quellenkritik und Hermeneutik bildet einen zentralen Baustein der NFDI.

4Memory verfolgt zwei Ziele. Erstens erschließt das Konsortium auf der Basis zunehmend digitalisierter und digitaler Daten sowie digitaler Methoden durch optimierte RDM-Infrastrukturen und Dienste **völlig neue Bereiche für die historische Forschung**. Das bildet die Voraussetzungen dafür, dass eine digital heterogene historische Forschung sich zukunftsfähig erneuert. 4Memory identifiziert Schlüsselinnovationen und fördert ihre Entwicklung, um sie in bestehende Strukturen und Prozesse zu integrieren und ihre breite Akzeptanz sicherzustellen. Unser zweites Ziel ist es, **eine methodologische Lücke zu verhindern**, indem wir gezielt digitale und analoge Methoden zusammenführen: Die digitale Transformation erfordert den Auf- und Ausbau von Kompetenzen, aber zugleich einen grundlegenden fachkulturellen Wandel. Das Arbeitsprogramm fördert daher digitale Kompetenzen in Universitäten, Archiven, Bibliotheken und Sammlungen und schafft grundlegende Strukturen und Anreize für innovative geschichtswissenschaftliche Forschung und Karrierewege.

Das Konsortium entwickelt ein **Portfolio von Schlüsseldiensten** und eine 4Memory-Forschungsdateninfrastruktur, die historische Wissenschaft und digitale Ressourcen systematisch miteinander verknüpft. Seine **Internationalisierungsstrategie** stellt sicher, dass globale RDM-Perspektiven in die nationalen Diskussionen und Verfahren einfließen, indem sie internationale Partner in Forschungs-, Gedächtnis- und Informationsinfrastruktureinrichtungen einbindet, und sie stärkt die existierende Zusammenarbeit in internationalen Strukturen und Leitungsgremien. Die Beratung in **rechtlich-ethischen Fragen** und ihre interdisziplinäre Diskussion sind ein zentrales Anliegen insbesondere hinsichtlich historisch geprägter gesellschaftlicher und globaler Asymmetrien. Durch die digitale Transformation **sichert NFDI4Memory die Bedeutung und kritische Funktion** der historischen Forschung in Gesellschaft, Kultur und Politik. 4Memory ist das Konsortium mit einzigartiger Expertise für historische Daten und Methoden und besitzt eine besondere Kompetenz, den gegenwärtigen Wandel der Wissensstrukturen zu reflektieren und damit die **Digitalisierung in eine kritische, langfristige Perspektive** zu stellen, die der NFDI als Ganzes dienen wird.

Applicant institution

Applicant institution	Location	Acronym
Leibniz Institute of European History	Mainz	IEG

Name of the Consortium Spokesperson

Spokesperson	Institution, location
Prof. Dr. Johannes Paulmann	Leibniz Institute of European History, Mainz

Co-applicant Institutions

Co-applicant institutions	Location	Acronym
Baden-Wuerttemberg State Archives	Stuttgart	LABW
Bavarian State Library	München	BSB
Deutsches Museum of Masterpieces of Science and Technology	München	DM
FIZ Karlsruhe – Leibniz Institute for Information Infrastructure GmbH	Karlsruhe	FIZ
German Historical Association (Verband der Historiker und Historikerinnen Deutschlands e.V.)	Frankfurt a. M.	VHD
Herder Institute for Historical Research on East Central Europe – Institute of the Leibniz-Association	Marburg	HI
Humboldt University of Berlin	Berlin	HU
Martin-Luther-University Halle-Wittenberg	Halle	MLU-HistData
Trier University	Trier	UT
University of Applied Sciences (h_da)	Darmstadt	h_da

Names of Co-spokespersons

Co-spokespersons	Institution, location	Task area(s)
Prof. Dr. Peter Haslinger	HI, Marburg	TA1
Prof. Dr. Helmuth Trischler	DM, München	TA1
Dr. Klaus Ceynowa	BSB, München	TA2
Dr. Katrin Moeller	MLU-HistData, Halle	TA2
Prof. Dr. Gerald Maier	LABW, Stuttgart	TA3
Matthias Razum	FIZ, Karlsruhe	TA3
Prof. Dr. Ursula Lehmkuhl	UT, Trier	TA4
Prof. Dr. Stefan Schmunk	h_da, Darmstadt	TA4
Prof. Dr. Torsten Hiltmann	HU, Berlin	TA5
Prof. Dr. Eva Schlotheuber	VHD, Frankfurt a. M.	TA5

Participants

Participating institutions	Location	Acronym
Akademie der Wissenschaften zu Göttingen Wissens-Aggregator Mittelalter und Frühe Neuzeit	Göttingen	WIAG
Arbeitskreis Provenienzforschung e.V.	Berlin	AK-PF
Archivschule Marburg – Hochschule für Archivwissenschaft	Marburg	ArchivSCH
Bayerische Akademie der Wissenschaften, München - Leibniz-Rechenzentrum - Kommission für bayerische Landesgeschichte	München	BAdW LRZ KBL

Bergische Universität Wuppertal - Institut für Grundlagenforschung zur Philosophiegeschichte - Interdisziplinäres Zentrum für Editions- und Dokumentwissenschaft (IZED) - Lehrstuhl für Digital Humanities	Wuppertal	BUW BUW-IGP BUW-IZED BUW-DH
berlinHistory e.V.	Berlin	BerHis
Bundesarchiv	Koblenz	BA
Clio-online e.V.	Berlin	Clio-online
Deutsche Nationalbibliothek - Coding da Vinci	Frankfurt a. M.	DNB CdV
Deutsches Schifffahrtsmuseum – Leibniz-Institut für Maritime Geschichte	Bremerhaven	DSM
digiCULT-Verbund eG	Kiel	digiCULT
Fernuniversität Hagen, Archiv des Instituts für Geschichte und Biographie	Hagen	FUH-Archiv
Forschungsstelle für Zeitgeschichte in Hamburg	Hamburg	FZH
Freie Universität Berlin - Lateinamerika-Institut - IRTG Temporalities of Future - Center für Digitale Systeme - Arbeitsbereich Geschichtsdidaktik - Universitätsbibliothek	Berlin	FU FU-LAI FU-IRTG FU-CeDiS FU-Didaktik FU-UB
Friedrich-Schiller-Universität Jena - Institut für Informatik - Historisches Institut	Jena	FSU FSU-II FSU-HI
Generaldirektion der Staatlichen Archive Bayerns	München	GDA
Georg-August-Universität Göttingen, Niedersächsische Staats- und Universitätsbibliothek Göttingen	Göttingen	SUB
Georg-Eckert-Institut – Leibniz-Institut für Internationale Schulbuchforschung	Braunschweig	GEI
Goethe-Universität Frankfurt, Universitätsbibliothek Frankfurt	Frankfurt a. M.	GU-UB
Germanisches Nationalmuseum – Leibniz-Forschungsmuseum für Kulturgeschichte	Nürnberg	GNM
Herzog-August-Bibliothek	Wolfenbüttel	HAB
Hessisches Landesarchiv	Wiesbaden	HLA
Institut für die Geschichte der deutschen Juden	Hamburg	IGDJ
Institut für Dokumentologie und Editorik e.V.	Köln	IDE
Institut für Sächsische Geschichte und Volkskunde e.V.	Dresden	ISGV
Institut für vergleichende Städtegeschichte, Westfälische Wilhelms-Universität Münster	Münster	IStG
Institut für Zeitgeschichte München−Berlin	München, Berlin	IFZ
Martin-Luther-Universität Halle-Wittenberg, Interdisziplinäres Zentrum für die Erforschung der Europäischen Aufklärung	Halle-Wittenberg	MLU-IZEA
Johannes Gutenberg-Universität Mainz, Zentrum für Datenverarbeitung	Mainz	ZDV
Konferenz für Geschichtsdidaktik e.V.	München	KGD
Leibniz-Institut für Bildungsforschung und Bildungsinformation, Bibliothek für Bildungsgeschichtliche Forschung	Berlin	DIPF BBF
Leibniz-Institut für Raumbezogene Sozialforschung	Erkner	IRS
Leibniz-Zentrum für Zeithistorische Forschung	Potsdam	ZZF
Ludwig Maximilian Universität München, IT-Gruppe Geisteswissenschaften LMU	München	LMU-ITGW

Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V., Max-Planck-Institut für Wissenschaftsgeschichte, Bibliothek	Berlin	MPIWG
Mediävistenverband e.V.	Frankfurt a. M.	MV
Mommsen-Gesellschaft e.V.	Freiburg	MG
Monumenta Germaniae Historica KdöR	München	MGH
Moses Mendelssohn Zentrum e.V.	Potsdam	MMZ
Otto Friedrich Universität Bamberg, Lehrstuhl für Medieninformatik	Bamberg	OFU
Rheinische Friedrich-Wilhelms-Universität Bonn, Arbeitsgemeinschaft Historische Grundwissenschaften / Nachwuchsnetzwerk Historische Grundwissenschaften	Bonn	AHiG
Ruhr-Universität Bochum, Centrum für Religionswissenschaftliche Studien	Bochum	CERES
Ruprecht-Karls-Universität Heidelberg - Centre for Asian and Transcultural Studies - Heidelberg Center for Transcultural Studies - Heidelberg School of Education - Research Council: Kulturelle Dynamiken in globalisierten Welten	Heidelberg	HDU HDU-CATS HDU-HCTS HDU-HSE HDU-RC
Sächsische Landesbibliothek – Staats- und Universitätsbibliothek	Dresden	SLUB
Stiftung Preußischer Kulturbesitz - Ibero-Amerikanisches Institut - Staatsbibliothek zu Berlin	Berlin	 IAI SBB-PK
Technische Hochschule Köln, Institute of Information Science	Köln	THK
Universität Bayreuth, Cluster of Excellence Africa Multiple	Bayreuth	ExClu AM
Universität Bielefeld, Fakultät für Geschichtswissenschaft, Philosophie und Theologie	Bielefeld	UBI
Universität Erfurt - Forschungsbibliothek Gotha - Professur für Globalgeschichte des 19. Jahrhunderts - Forschungszentrum Gotha	Erfurt	UE FB Gth UE-GG UE-FZG
Universität Hamburg - Zentrum für nachhaltiges Forschungsdatenmanagement - Cluster Understanding Written Artefacts	Hamburg	UHH UHH-FDM ExClu UWA
Universität Kiel, Universitätsbibliothek – Fachinformationsdienst Nordeuropa	Kiel	UK-UB
Universität Leipzig - Digital Humanities Lab - Global and European Studies Institute - Graduate School Global and Area Studies - Leipzig Research Center Global Dynamics - Lehrstuhl für Alte Geschichte, Historisches Seminar - Universitätsbibliothek Leipzig	Leipzig	UL UL-DHLab UL-GESI UL-GSGAS UL-RCG UL-AG UL-UB
Universität Mannheim, Universitätsbibliothek	Mannheim	UBM
Universität Marburg, Deutsches Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg	Marburg	DDK
Universität Passau, Lehrstuhl für Digital Humanities	Passau	UP-DH
Universität Regensburg, Gesellschaft für Sozial- und Wirtschaftsgeschichte, Lehrstuhl für Wirtschafts- und Sozialgeschichte	Regensburg	GSWG
Universität Rostock, Institut für Medienforschung	Rostock	UR
Universität Tübingen, Zentrum für Datenverarbeitung (ZDV-T)	Tübingen	ZDV-T

Universität zu Köln - Cologne Center for eHumanities - Data Center for the Humanities - Lehrstuhl für die Geschichte der Frühen Neuzeit - Zentrum für LehrerInnenbildung	Köln	UzK UzK-CCeH UzK-DCH UzK-FNZ UzK-ZfL
Verband deutscher Archivarinnen und Archivare e.V.	Fulda	VdA
Verbundzentrale des GBV (VZG), Göttingen	Göttingen	VZG
Verein für Computergenealogie e.V. (CompGen), Köln	Köln	CompGen
Westfälische Wilhelms-Universität Münster Wirtschaftshistorischer Ausschuss, Verein für Socialpolitik	Münster	WA-VfS
ZBW – Leibniz-Informationszentrum Wirtschaft	Kiel	ZBW

Contributions of Participants

AHiG: uses its network and communication channels to debate data culture; contributes to related events, especially aimed at early-career researchers. (TA5)

AK-PF: expertise on provenance, history of collections and institutions, processes of authentication, legal and ethical questions (related to research data). (TA1, TA5)

ArchivSCH: expertise on archival description, archival law, ethics and digital archiving; intermediary for conference events. (TA1, 2, 4, 5)

BA: mandate to secure archival material in the long term and to make them available for use; supports the efforts of harmonizing rules and standards for documentation. (TA2)

BAdW: develops and promotes recommendations on ontologies and multilingualism. (TA2)

BerHis: advice on data security, ethics and quality assessment; community involvement; production, dissemination and (re)use of historical data. (TA5)

BUW: philosophical-historical expertise, development of data criticism and sustainable RDM; experience in digital methods; development of a professional data culture. (TA4, TA5)

CERES: intermediary to religious studies; methods training and outreach; Buddhist material culture in South and West Asia; historically oriented research in religions and art. (TA4)

Clio-online: responsible for TA5-M5; editorial system of quality control and review for research data (publications); critical peer discourse (with H-Soz-Kult); Clio-Guides. (TA5)

CompGen: contributes software-based services for collaborative work and Linked Open Data (DES, GOV), network of historically engaged citizen science. (TA2, TA3)

DDK: expertise on data quality regarding data and metadata standards, contributes to best-practice guidelines; expertise on and provision of authority data. (TA1, TA2)

digiCULT: expertise on authority data, metadata, cataloguing, documentation and description of research data; key services, expertise for knowledge graph. (TA2, TA3)

DIPF / BBF: expertise on subject-specific applications, historical classifications, ontology-based modelling, RDM; FID Educational Science and Educational Research. (TA2, TA5)

DNB: data-quality requirements for digital information infrastructures; intermediary in quality assurance initiatives; GND / interdisciplinary authority data; knowledge graphs. (TA2, TA3)

DSM: quality control for research data and collections; network for researchers implementing data-quality standards; hub for expertise and digital infrastructures. (TA1, TA5)

ExCIu AM: designing knowledge graph and virtual indexing layer; contributes non-European repositories; diversity of regional, legal, and epistemological dimensions of RDM and data criticism; perspectives on Global South. (TA1, TA3, TA4, TA5)

FSU: expertise in digital research and teaching; standards for 3D and XR data; implementation of LOD; medieval data to the Knowledge Graph; multi-modal machine learning. (TA2, TA4)

FU: expertise in area-specific data literacy, multilingualism (Latin America, East Asia); e-learning methods; Oral history; key service (Oral-History.Digital). (TA2, TA3, TA4)

FUH-Archiv: expert in oral history; provider of historical interview collection, expertise on data extraction; creation of historical authority data and curation tools. (TA2)

FZH: Oral History Archive; Research Centre for Contemporary History with expertise on interviews as research data; research-ethics and secondary analysis. (TA5)

GDA: archival expertise; quality requirements for archival data; targeted courses: pupils, students, prospective archivists; infrastructure for workshops, conferences. (TA1, TA2, TA4)

GEI: International textbook research; onsite and digital access to international educational media and curricula; digital tools for data analysis. (TA1, TA5)

GNM: expertise on metadata, cataloguing, documentation and description of research files; contributes to 4Memory ontology and metadata schema. (TA2)

GSWG: RDM of economic and social history; expert in development of authority data and indexing tools; intermediary for dissemination of data standards. (TA2)

GU-UB: advice on multilingualism and ethical issues; expert on authority data and data curation; data provider and intermediary. (TA2)

HAB: expertise on metadata, cataloguing, documentation and description of research data; engages in Data Editorial Framework and FAIR Data Fellowship group. (TA1, TA2, TA5)

HDU: expertise in Global and Area Studies, Asia in a Global Context; contributes digital tools, methods of reading, digital didactics; course programs in digital literacy. (TA4)

HLA: key service for digital preservation (DIMAG), related expertise and advisory services; contributes experience and knowledge in handwriting recognition. (TA3)

IAI: area-specific expertise from social sciences and humanities dealing with Latin America, the Caribbean, Spain and Portugal; librarian and cultural education. (TA4)

IDE: publishers of review journal RIDE; criteria for evaluating digital data and scholarship; expertise in evaluating digital research data. (TA5)

IFZ: authority data and data documentations; integrating data in bibliographies; knowledge graph input; key service (research data repository); EHRI-Portal for Data Space. (TA3)

IGDJ: expert and provider of collections of sources, texts and images on Jewish history; inter-institutional thesaurus and various research tools for this subject area. (TA2)

IRS: expertise and networking in socio-spatial research and RDM; intermediary of historical and social sciences; dissemination. (TA5)

ISGV: Digital Historical Gazetteer of Saxony and biographical electronic database; referencing data with Wikidata, BEACON and GND; scaling models for standards data. (TA2)

IStG: technical and conceptual expertise on geographical authority data and controlled vocabularies; provides geodata to the consortium. (TA2, TA3)

KBL: expertise on documentation practices, bibliographies, ontologies, canonical URIs. (TA2)

KGD: training methods qualifying future history teachers; historical data criticism; intermediary to the community of history teachers; dissemination of content recommendations. (TA4)

LMU-ITGW: networking with RDA and Go Fair; expertise on metadata exploitation and research data documentation; contribution to guidelines. (TA1, TA2)

LRZ: software development expertise with a focus on FAIRification of data collections. (TA3)

MG: surveying needs from the classics community; expertise on digital transformation and the use of research data; communicates discourses on data-driven research. (TA5)

MGH: expertise in authority data working group; contribution to Metadata Competence Unit, focusing on the topic of canonical URIs and their use. (TA2)

MLU-IZEA: provides digital editions of letters and newspapers; expertise on enrichment of data and Semantic Web formats / techniques. (TA2)

MMZ: Thesaurus on Jewish History; coordinates development of digital resources; member of International Authority Data Expert Group. (TA2)

MPIWG: contributes to 4Memory metadata schema; multilingualism and connectivity of the data space via APIs; design of the knowledge graph and virtual indexing layer. (TA2, TA3)

MV: bidirectional communication to its large medievalist community; expert and host for events related to research data for all disciplines concerned with the Middle Ages. (TA5)

OFU: working groups on authority data, data quality and metadata harmonisation; contributes its Data Modeling Environment and expertise to creating the Data Space. (TA2, TA3)

SBB-PK: register of standards data; provides authority data; advises on services, multilingualism, metadata standards, technical interfaces and data formats; design of Data Space and virtual indexing layer; communication measures. (TA2, TA3, TA5)

SLUB: register of standards data; provider of authority data, geographic information system services; contributes to virtual indexing layer. (TA2, TA3)

SUB: comprehensive metadata expertise; surveying data quality requirements; community discourse of the role of research data in the historically oriented humanities. (TA1, TA5)

THK: expertise in RDM, information management, digital literacy (data and information literacy), e-science and information management. (TA1)

UBI: digital methods and data-driven research in the humanities; development of a new data culture; Roadshow on research data practices. (TA5)

UBM: expertise in RDM of economic and social history data; service for hosting data; expert for subject-specific authority data and data extraction. (TA2)

UE: expertise in creating digital infrastructures; key service (FactGrid); machine-interpretable knowledge representation; data criticism and RDM. (TA2, TA3, TA4, TA5)

UHH: expertise in building databases, data repositories and APIs; technical requirements for linking data; creating sustainable RDM solutions. (TA2, TA3, TA4)

UK-UB: as a bridge to Northern European information infrastructures and memory institutions, contributes its expertise on metadata harmonisation on an international scale. (TA2)

UL: expertise in RDM, digital tools in area and global studies, human geography and cultural studies; develops MA-modules on digital research culture; use case “spatial semantics”; experience on APIs and viewer technologies. (TA2, TA3, TA4, TA5)

UP-DH: provides expertise in digital history methods and epistemological questions; advises on young researcher networks. (TA5)

UR: expertise GIS systems, geodata and historical socio-demographic mass data; occupational taxonomies, authority data and data curation. (TA2)

UzK: evaluating and developing standards for data description and documentation; engages in data editorial framework; contributes to analysis of data literacy. (TA1, TA2, TA4)

VdA: professional training capacities for archives in the fields of long-term archiving, authority data and data standards; intermediary to the community of archives. (TA2, TA3)

VZG: provider of key service (DANTE), building the knowledge graph, integrates collections into Data Space, expertise in conceptualising a virtual indexing layer. (TA3)

WA-VfS: expertise on and data from qualitative and quantitative research in economic and social history; corresponding data series; expertise on curation of time-variable data. (TA2)

WIAG: medieval and early-modern data collection project: assists in creating a historical authority data registry, provides data and expertise on authority data. (TA2)

ZBW: provider of “Standard Thesaurus for Economics”; contributes to the international authority data expert group for the development of historical vocabularies. (TA2)

ZDV: technical coordination; expertise on information infrastructures and architectures; RDM services; technical cross-cutting topics; project infrastructure. (TA3, TA6)

ZDV-T: provider of technical infrastructure for the 4Memory Data Lab; expertise machine learning (TA3)

ZZF: expertise on digital history, contemporary history, digital publication platforms, oral history interviews and audio-visual research data, authority data, RDM strategies. (TA2, TA5)

Participating individuals

Name(s)	Institution, location	Acronym
Prof. Dr. Andreas Kuczera	Akademie der Wissenschaften und der Literatur Mainz / Technische Hochschule Mittelhessen	AKucz/AdW
Prof. Dr. Christina von Hodenberg	Deutsches Historisches Institut London	DHI London
Dr. Sandra Dahlke	Deutsches Historisches Institut Moskau	DHI Moskau
Prof. Dr. Thomas Maissen	Deutsches Historisches Institut Paris	DHI Paris
Prof. Dr. Martin Baumeister	Deutsches Historisches Institut Rom	DHI Rom
Prof. Dr. Simone Lässig	Deutsches Historisches Institut Washington	DHI Washington
Prof. Dr. Franz Waldenberger	Deutsches Institut für Japanstudien, Tokyo	DIJ Tokyo
Dr. Sebastian Schwecke	Max Weber Forum für Südasienstudien, Delhi	MWF Delhi
Prof. Dr. Birgit Schäbler	Orient Institut Beirut	OIB
Prof. Dr. Rüdiger Hohls	Institut für Geschichtswissenschaften, Humboldt-Universität zu Berlin	RHohls/HU

Contributions of Participants

AKucz/AdW: provision of a use case, supplying a medieval data corpus (Regesta Imperii) for an interdisciplinary authority data working group. (TA2)

DHI London: subject-specific vocabularies; training materials, consulting on digital literacy; digital transformation in historical research, communicative measures. (TA2, TA4, TA5)

DHI Moskau: intermediary with German, Russian and international communities; multilingual data; international perspectives on data literacy, legal/ethical issues. (TA1, TA2, TA4, TA5)

DHI Paris: intermediary with French historically oriented humanities; authority data; international RDM perspective; digital history training for early career scholars. (TA2, TA5)

DHI Rom: provision of data, metadata and authority data; developing models, use cases, workflows, tools and services; contributes to a DL profile; digital methods. (TA2, TA4, TA5)

DHI Washington: training material for data literacy; international network-building; use case on multilinguality; digital methods in North America. (TA4, TA5)

DIJ Tokyo: training material for data literacy; expert groups on digital transformation and multilinguality; international network-building; digital methods. (TA4, TA5)

MWF Delhi: Indian-based data collections; non-Western perspectives/training materials on data literacy; developing data culture; internationalisation. (TA1, TA4, TA5, TA6)

OIB: expertise on multilinguality; surveying current state; recommendations on data quality; non-European perspectives in historical data criticism. (TA1, TA4)

RHohIs/HU: extensive expertise on scientific communication, data-based research and academic networks in history and digital humanities. (TA5, TA6)

DFG review boards

DFG review boards (DFG-Fachkollegien) reflecting the subject orientation of 4Memory:

Review Board	Subject areas
101 Ancient Cultures	101-03 Ancient History
102 History	102-01 Medieval History 102-02 Early Modern History 102-03 Modern and Current History 102-04 History of Science
106 Social and Cultural Anthropology, Non-European Cultures, Jewish Studies and Religious Studies	106-02 Asian Studies 106-03 African, American and Oceania Studies 106-04 Islamic Studies, Arabian Studies, Semitic Studies 106-05 Religious Studies and Jewish Studies
108 Philosophy	108-01 History of Philosophy
109 Educational Research	109-01 General Education and History of Education
112 Economics	112-05 Economic and Social History
Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (AWBI)	Archival studies Library studies Museum studies

2 Scope and Objectives

NFDI4Memory will set the pace for institutional, cultural and methodological innovation in historical research. 4Memory for the first time systematically **integrates research, memory[1] and information infrastructure institutions** into a common digital research data infrastructure, enabling a **deep, sustainable transformation** of the historically oriented humanities. With its unique expertise in digital historical source criticism and hermeneutics, it drives forward key elements of the future knowledge order for the NFDI. It offers **innovative products and processes** enabling cultural change and sustainable RDM methods and services. The consortium unites **established, pioneering co-applicants** with vast experience in digital methods with over 70 participants from our community. The applicants represent the disciplinary and epochal breadth of historical research through the German Historical Association (one of the largest humanities associations in Europe with more than 3,400 members), top German universities and excellent research institutes from the Leibniz Association. The consortium's memory institutions are key players in RDM for historical research, including state archives like the Baden-Wuerttemberg State Archives and the German Federal Archives as well as the national training college for archivists in Marburg. The Bavarian State Library together with Specialist Information Services run by other state and university libraries across Germany – alongside collections such as the Deutsches Museum – are major contributors to RDM standards and methods nationally and internationally. Our co-applicant FIZ Karlsruhe's technical and information science experience benefits 4Memory's service portfolio; its extensive involvement with the NFDI supports our contributions to cross-cutting topics.

4Memory's combination of research, memory and infrastructure institutions ensures both effective leadership and a wide reach for measures to accelerate the historical disciplines' digital transformation. Our revised application strengthens our emphasis on **innovative methods** while maintaining one of our greatest strengths: our deep embeddedness in our heterogeneous community. We have two connected aims. First, the growth of born-digital data and the development of cutting-edge digital methods means that **entirely new areas of historical research are being opened**, requiring improved RDM infrastructures and digital services. 4Memory will contribute to pushing forward innovation in historical research by providing RDM guidelines and digital services for a research community still at an early stage of digital transformation: for example, only a tiny proportion – as little as two per cent – of archival holdings have been digitised, and analogue sources remain central to historical study. This distinguishes 4Memory from other consortia and points to our significance in transforming a quantitatively large group of scholars in the humanities. Thus, our second aim is to **bridge a methodological gap**

[1] Memory institutions include archives, libraries, museums, memorial sites and heritage institutions as well as any provider of digital collections serving as memories of mankind. The term provides an overlap with the GLAM acronym (galleries, libraries, archives, and museums) but extends its scope and provides a focus on the concept of memory, which is why we use "memory institutions" here.

by bringing digital and analogue methodologies together. Given the breadth of our community's involvement (a point praised in our first application), our measures promise to have a major impact on innovating historical scholarship. Historical perspectives are indispensable to the NFDI's mission and the digital humanities. In this preface, we demonstrate our key role, showing how we address suggestions for improvement from the review process.

1. Measures to foster and promote innovation in our community

The humanities are undergoing rapid and dynamic development through advancements in digital technologies and methods, both those driven by the needs of working with increasing amounts of digitised sources and born-digital data and by continuing improvements in the digitisation, provision and analysis of data. We aim to identify key innovations, evaluate their value to our community, foster their development, integrate them into existing structures and processes and promote their wider adoption. 4Memory will both **encourage the development of advanced RDM strategies at the cutting edge of digital history** and **set up equally innovative procedures that will bring the wider historically oriented community forward**. We see ourselves as pacesetters for the digital transformation and as enablers of a wider disciplinary advancement.

Both visions shape our extended RDM Strategy and are embodied in our work programme. The **"Innovation framework"** (TA6-M3) encourages advanced methods, encompassing tasks to facilitate and disseminate innovations. The "**Incubator for complementary and innovative actions"** (TA6-M3-T1) will enable us to move forward by dynamically integrating pathbreaking new developments into 4Memory's work programme. Running parallel to that task, the **"Participant and partner liaison"** (TA6-M3-T2) ensures a trustful, direct and continuous knowledge exchange on innovative ideas and up-to-date developments. We also introduce a new format for innovative and experimental approaches: the 4Memory Labs. The **4Memory Data Lab** (TA3-M5) offers an *infrastructure* setup allowing experimental approaches to explore and analyse datasets in the 4Memory Data Space. The **4Memory Methods Lab** (TA5-M2) seeks to identify, evaluate and promote new *data-driven methods* in the historically oriented disciplines. The **4Memory Data Literacy Training Lab** (TA4-M1) contributes to transferring *expert knowledge* into academic programmes, school curricula and state educational policies.

4Memory's portfolio ensures the **sustainability of innovative procedures and products**. Setting guidelines for data quality (TA1), mainstreaming data literacy education (TA4) and ensuring metadata interoperability through standardisation and harmonisation (TA2) are integral to fostering innovation. Participating institutions provide key services (e.g. those related to long-term hosting) as part of their institutional activities and can thus adopt innovations from the community throughout and beyond the funding periods. In addition, international knowledge exchange in digital history via our partnerships allows us to identify new developments from the global community while promoting and integrating our own innovative ideas.

2. Strategies for developing and expanding digital competencies

Digital transformation requires capacity building and cultural change, aims we pursue throughout our consortium, e.g. by enhancing digital competencies for those working in archives and libraries (TA1, TA2). In two complementary task areas 4Memory ensures the widest possible development of digital competencies (TA4) and the creation of incentives and structures to make them a priority (TA5).

Task Area 4 (Data Literacy) brings together university departments, professional associations for historical didactics, archives and archival training colleges to develop **cutting-edge teaching concepts and training materials** at all levels of history education. It provides services for **capacity building**, including "teach the teachers" webinars, and ensures existing DL curricula are for the first time **systematically networked** across Germany's diverse, federalised education system. To impact curricula we develop and implement **target-specific modules** and a **model curriculum** for BA and MA study programmes (TA4-M1-T2); produce **period-specific syllabi** as best-practice solutions for teaching historical data criticism and sustainable RDM (TA4-M2-T3); organise **DL-Labs** for junior and senior scholars focusing on area-specific RDM and addressing the legal and ethical challenges of digital area histories (TA4-M3-T2); host **webinars** for the most needed digital methods, tools, and techniques (TA4-M4-T2); set-up **regional training centres** (TA4-M4-T3); create a **virtual counselling hub** (TA4-M4-T4); organise a **DL-training accreditation service** (TA4-M1-T4); and pursue an **advocacy and outreach strategy** addressing the responsible public agencies (TA4-M1-T4).

Cultural change means more than new skills: **Task Area 5 (Data Culture)** ensures RDM and digital methods become valued in our disciplines. Those already engaging in pioneering digital methods in historical fields must receive **professional recognition** for their efforts, something current frameworks insufficiently provide. TA5's revised design aims to ensure that both a) **incentives** are in place (particularly for early career researchers) to encourage the development of digital skills and the improved handling of research data and to recognise such efforts and b) **processes** (e.g., of **evaluation and review,** TA5-M4) are established to ensure a new data culture emerges. We must **understand the implications** of changes in the production of historical knowledge and the consequences of data-driven methods (TA5-M1). In particular, artificial intelligence and machine-learning procedures will cause a long-term transformation that needs to be discussed, evaluated and critically integrated. 4Memory has the **pathbreaking** task to establish new **platforms** and **procedures** and set advanced examples (TA5-M5).

3. Internationalisation, with special reference to harmonisation of metadata

Advancing RDM nationally means engaging with RDM *inter*nationally to avoid patchwork solutions, benefit from leading advancements and share our own innovations worldwide. Our carefully integrated internationalisation strategy (coordinated in TA6-M2-T4) will ensure that global perspectives on RDM are brought into the defining discussions and procedures; unite international partners in research, memory and information infrastructure institutions in a new degree of cooperation and dynamically expand their existing networks; and strengthen essential cooperation with transnational structures and governing bodies. Our strategic participants include **German-based institutions with international research focuses** and extensive contacts in specific regions, e.g. the Excellence Cluster Africa Multiple at the University of Bayreuth, the Global and European Studies Institute at the University of Leipzig and the Heidelberg Center for Transcultural Studies. Among our participants are the **German Historical Institutes located abroad,** namely those in London, Rome, Moscow, Washington, Beirut, Delhi and Tokyo, each with their regional networks and special areas of knowledge. Our memory and infrastructure institutions have extensive experience in **international RDM networks and initiatives** (e.g. Research Data Alliance, Go Fair Initiative, DARIAH-EU). We also have strategic **partnerships with institutions and organisations throughout Europe and beyond**, such as the Huygens Institute, the Roy Rosenzweig Center and the Data for History consortium in the field of digital history.

With regard to specific internationalisation topics, substantial **cooperation will take place within the Task Areas**. The harmonisation of **metadata** has an important role in this context. The consortium pursues a differentiated and discipline-based concept of metadata, authority data, taxonomies and ontologies (TA1-M4; TA2-M1, M2, M4, M5). Ensuring international compatibility is a key part of our service offerings, such as our Data Space (TA3-M3). Through 4Memory's international network we will spark a proactive knowledge exchange regarding metadata and develop cooperation in international working groups and committees. Here, 4Memory profits from the long-term participation of its applicant institutions in relevant **international committees** (CIDOC, LIBER, CERL, Europeana working groups etc.) and their experience in developing and using **international metadata standards** (MARC/RDA, EAD, LIDO, DataCite, METS/MODS), technical standards on interoperability (IIIF) and authority data (GND, VIAF, Getty, HISCO, ISCO/KldB, GOV). Our particular contribution lies in ensuring issues of historical research – e.g. temporality, changing terms, provenance and contextualisation – as well as ethical questions in data and metadata receive due attention. The experiences, knowledge and connections that 4Memory gains through these activities will flow **into the NFDI** via our contributions to its cross-cutting section "Metadata, Terminologies, Provenance".

4. Ethical and legal issues

Historical RDM raises a host of ethical and legal issues. **To whom do digital historical data belong**, i.e. which groups participate in the discussions had and decisions made about them? **Ease of access to data across cultural boundaries** means that data can be more easily used in contexts foreign to their creation, raising questions of cross-cultural responsibility and fairness. **Machine learning / artificial intelligence** methods pose complex problems, not least their unreflected use, loss of original contextualisation and potential deliberate misuse. The "**right to be remembered**" must be considered: using historical data is always a selective process and we need to consider which data are taken up into the historical record and which are left out. **Maintaining privacy in contemporary historical sources** is vital. Some periods of historical work concern people who are still living (and their immediate descendents): what are the legal and ethical boundaries in using such sources? Contemporary history (much like anthropology, social sciences and medicine) often relies on interviews that are recorded (whether as audio or video) and records from certain kinds of state and welfare institutions (such as prisons and hospitals) that may need **anonymisation**. What ethical questions are raised by data created in **contexts of injustice**, such as authoritarian regimes or under colonial rule? The ethical question to what extent data – and publications based upon them – *should* be "**open access**" is inseparable from the legal frameworks in which data and research results *are* made available. We must also ensure that access to quality data is not needlessly hindered by expense. While for many historians, the problem of sources and data being in **copyright** is a secondary issue (because they address periods in the distant past), those dealing with 20th and 21st century topics often face serious limitations in using – or even referencing – recent materials. Researchers require guidance regarding these questions.

We pursue several **strategies**: assessing community needs, providing forums for discussion and advice and engaging with professional associations and political decision-making bodies to develop solutions (TA1-M2; TA4-M1, M3, M4; TA5-M5). The key role of **memory institutions** in 4Memory makes us ideally placed to address the ethical and legal issues of archival sources. Our participants include **experts in oral sources and contemporary history**. We also work closely with **other consortia** specialising in areas relevant to historians: for example, our MoU partners in NFDI4Culture have a focus on copyright issues; KonsortSWD shares our interest in the legal issues pertaining to interviews, institutional data and sociological big data. 4Memory will, furthermore, create a "**Special Committee on Research Ethics**" (based in TA6-M2-T4 with contributions from all TAs) composed of experts on both the Global North and the Global South to discuss challenges and to advise us regarding solutions. The synergies resulting from these diverse, interlocking activities will enable us to contribute to developing the **legal and ethical topics in the NFDI**, such as via the cross-cutting section "Ethical & Legal Aspects".

5. Our value to the NFDI

All disciplines require contextualisation and source criticism; 4Memory's key contribution is our community's expertise in historical and hermeneutical approaches: we are the only proposed consortium focused on historical data. Data criticism means considering the integrity and authenticity of data: historians go further, reflecting fundamentally upon the **historicity of data**, including technical changes in data storage and formats. Research data and the sources they're based on are not only heterogeneous, extending from texts, statistics and images to audiovisual media and born-digital information (in contemporary history). We must **understand their origins in historical contexts**. Interpreting them requires knowing where, how and in which (digital) format they were produced. This is equally true for interdisciplinary research, e.g. with data on the long-term modelling of climate change or infectious diseases. The empirical basis of historical knowledge – our sources – were created amid diverse social and cultural relationships, in which, e.g. the meaning of names and place names – and even of terms, categories and concepts – were different and in **a process of change**. The provenance of digital historical sources or documents must, as with their analogue counterparts, be transparent and comprehensible. Knowledge about the data's historicity must not be lost through digital representation and can be made more accessible and even improved. We are well placed to reflect upon the history of knowledge structures and to place **digitalisation in a critical, long-term perspective**. We will bring these unique contributions into the NFDI as data – irrespective of disciplinary perspective – are fragile, incomplete and emerge in dynamic relations. We thus add substantial value to the research data infrastructures, to the benefit of our community and to the NFDI as a whole.

4Memory will contribute to the NFDI through its engagement in NFDI structures, both **as a consortium** (our work programme earmarks resources for coordinating such activities, TA6-M3-T3) and through its **individual members** (both co-applicants and participants). Most of our co-applicants, and many of our participants, are already institutional members of NFDI e.V. Our consortium's unique combination of institutions will enable valuable contributions from different perspectives: those of archives and libraries, collections, historians and citizen science. The NFDI will also benefit from the participation in our consortium of institutions experienced in political advocacy and public engagement, particular legal issues (e.g. archival law) and of the professional society (the German Historical Association) that represents historians at the university and secondary school levels as well as those working at memorial sites.

Our substantial base of participants enable connections to other NFDI consortia, with whom they are also involved. We have made individual agreements with other NFDI consortia to cooperate on specific aspects of our work programme (see 3.2). We are co-authors of a **Memorandum of Understanding** with three fellow humanities consortia: 4Culture, Text+ and 4Objects. Meeting regularly, we have already been coordinating our joint contributions to multiple cross-cutting sections. For more details see section 3.2.

2.1 Research domains or research methods addressed by the consortium, specific aim(s)

NFDI4Memory serves **all humanities disciplines that make significant use of source materials produced in or describing the past, thus requiring historical contextualisation and critical analysis**. It focuses on the historical discipline as such – from ancient and medieval history to early modern, modern and contemporary history and covers Europe as well as world regions and cultures beyond – and incorporates other disciplines relying on historical data, such as religious studies, area studies and the history of philosophy. From an institutional perspective, **archives, libraries, collections, and museums** are integral parts of historical research, both as places of digitisation and as partners in the research process; **4Memory thus aims to transform the existing cooperation based on a traditional division of labour into a digitally integrated relationship between historical research and memory institutions with regard to research data**. In this process, 4Memory will establish new forms of cooperation among research and memory institutions, computing centres and digital service providers.

Those working historically use strategies of source criticism – the critical questioning of the origins and contexts of historical sources centring on questions such as authorship, intention, audience and authenticity – that have long been applied to analogue data; however, we are still at the start of creating forms of **"digital source criticism" for digitised material and other new kinds of historical data** (Fickers 2012, Föhr 2017, Lehmkuhl 2020). In historically oriented disciplines "research data" include all data collected, described, analysed, evaluated, generated and stored in machine-readable form for maintaining the traceability of research results as well as for archiving, citation, contextualisation and further processing. Among this data one finds, e.g., texts (handwritten or printed), images, maps, objects, statistical databases, audio recordings and films from across the world. There is a growing amount of "born-digital" data, even if historians still mainly work with digitisations of analogue materials. The interface between digital data and analogue material is central to our fields. Different types of data and research methods require modular, flexible approaches; **we aim to ensure researchers can access up-to-date methods, services and advice and to advance purely digital research, supporting a transformation of the historical disciplines.**

Beyond technical issues of data formats and standards – many common to other fields – historical data requires the **contextualisation** at the heart of historical scholarship. Historical data are not only heterogeneous regarding where, how and in what format they were produced but also in terms of when and by whom they were created and the reasons for, and contexts of, their preservation. The meanings of words, the names of places and even the underlying categories of knowledge have changed – often substantially – over time. Accounting for space, time and historical actors is key to studying history, which integrates multiple perspectives, relationships and forms of ambiguity. These technical and conceptual factors must be made transparent and re-traceable when presenting historical data; also, the specificities of the contexts in which data

were created in the past and then collected, and potentially transformed, in the (ever advancing) present need to be kept in mind and on the digital record. **4Memory offers expertise in historical methods and hermeneutical approaches to the NFDI, supporting the contextualisation and source criticism of data vital to all disciplines**.

With digitalisation, the borders between scholarly and non-scholarly knowledge are often blurring. People's relationship to knowledge changes when scholarship is no longer found only in university libraries in printed monographs, edited collections or journal articles but rather on any screen, where it is hard to differentiate it from questionable information that, on the surface, may look exactly the same. Current events make clear that **the humanities – particularly the historically oriented disciplines – must maintain their socially critical functions**.

2.2 Objectives and measuring success

4Memory's specific objectives can be summarised under the acronym "**LINKAGE**", providing a community-specific agenda for meeting the NFDI's goals.

1. Linking Research, Memory Institutions and Infrastructures: 4Memory for the first time systematically integrates research, memory and information infrastructure institutions into a common digital research data infrastructure by establishing sustainable links between them.

2. Integrating Historical Source Criticism into Data Services: The conditions created by digitised and born-digital sources require adapting established traditions of source criticism to new research realities. This is a core aspect of historical research, and we will ensure our community's expertise in hermeneutical approaches also contributes to other disciplines.

3. Network of Historically Oriented Research Communities: 4Memory connects the historically oriented humanities, above all the large community of historians and including many fields that work with historical data. We enable interdisciplinary exchanges with the social sciences and cultural studies and reflection about non-European data.

4. Knowledge Order for the Digital Future of the Past: Knowledge is produced, classified and disseminated in systems that create and connect institutional and disciplinary knowledge spaces. Transforming the humanities' existing knowledge order from one largely based on printed material and traditional systems of archival and library classification to a digital one requires collaborative processes, new criteria for "verified" knowledge and innovative procedures for its evaluation and citation.

5. Advancing the Analogue / Digital Interface of Historical Source Material and Data: 4Memory will encourage the development of advanced RDM strategies at the cutting edge of digital history. Since digital media coexist with analogue media and historians and archives will always also rely upon physical sources, we will bring digital and analogue methodologies together, making more data available for innovative analysis.

6. Generating Standards for Historical Research Data and Sustainability: Maintaining research-data quality requires discipline-specific, cross-disciplinary and international data-handling standards. These priorities must be fully integrated into the research process, requiring new research data management personnel as well as empowering historians themselves to ensure their data meets standards of quality, transparency and reusability.

7. Education and Citizen Participation: Improving data literacy is essential to the future data culture in our community at all levels of the educational system. In a time when research is faced with proliferating sources of information of vastly different, but not clearly distinguishable, levels of reliability, empowering citizens to use digital information thoughtfully is a must.

4Memory will improve historical RDM by linking those at its constantly advancing cutting edge with those at an early stage of digitalisation. Details of criteria, measurements and evaluation procedures for quality assurance are given in our TAs' work programmes, along with milestones and deliverables. Our efforts will be successful to the extent that we achieve the following:

- Ensuring the quality of historical data by establishing clear **guidelines for dealing with data and metadata**;
- Increasing the interoperability and reusability of historical data and data collections – both within and beyond the historically oriented disciplines – **through the harmonisation and implementation of well-defined standards**;
- Developing the interconnectivity of existing data collections (analogue and digital) through **expanding the ability to search across platforms** and by **making varieties of historical data increasingly available** for research purposes;
- Expanding the **visibility and usability of research data**, also for an interested public of citizen scientists, political actors and social activists;
- **Strengthening national and international networks and partnerships** in RDM, primarily for the historically oriented humanities but with a view to cross-disciplinary and NFDI-wide cooperation;
- **Improving the digital literacy** of historical scholars regarding RDM, enabling them to work more effectively with their partners in memory institutions and empowering them to independently employ data management best practices;
- Raising awareness for the need to ensure that data are handled in accordance with **legal and ethical standards** and providing forums for the discussion and development of relevant recommendations;
- **Encouraging a cultural change** in historically oriented fields through an active dialogue with the relevant communities to provide incentives for researchers to improve their handling and understanding of research data, **increase the status of research data management and drive forward the development of digital methods in research**.

3 Consortium

3.1 Composition of the consortium and its embedding in the community of interest

4Memory brings together historical research, memory and information infrastructure institutions. The co-applicant institutions have extensive RDM experience in the following areas:

- The provision, preparation and long-term preservation of research data – as well as indexing, cataloguing, publishing them and enabling their future reuse – are core activities of the participating archives, libraries, museums, collections, and information infrastructure institutions;
- The development of workflows and services tailored to historical research are key competence areas of the involved university service centres, research institutes and information infrastructures;
- In modelling the historical domain, the co-applicant data centres, libraries and departments of information science pool their respective expertise on internationally standardised metadata and linkage concepts;
- Training in data literacy and the development of digital methods are areas of expertise for our participating digital history departments and DH Labs;
- Initiatives on promoting and developing innovative digital methods and data culture – and a decades' long experience with our community – have been developed by working groups of the relevant professional societies of historians and archivists.

3.1.1 Participating institutions

The Leibniz Institute of European History (IEG, applicant institution): The IEG in Mainz is an independent research institute with a global scholarly network and a substantial information infrastructure. A member of the Leibniz Association, it has received grants from bodies such as the German Research Foundation (DFG), Federal Ministry of Education and Research (BMBF), Volkswagen Foundation, and the European Commission (Horizon 2020). It has a broad focus on the modern history of European societies; transnational relations within and beyond the continent; and religious history, including Judaism and Islam. Its residential scholarship programme has benefitted some 1,500 scholars. Its library offers 90,000 printed titles and 900,000 licensed online resources. It is a pioneer in digital history and innovative open-access publishing, providing high-value resources such as European History Online (EGO) and IEG Maps. Since 2019 the Digital Historical Research Unit | DH Lab has shaped the transformation of historical scholarship. It carries out projects, develops software and DH methods, curates collections and designs digital-literacy concepts. The IEG is part of regional DH networks (mainzed), several national and international networks (DARIAH), and consortia on digital research infrastructures including ESFRI projects (RESILIENCE). It is a member of NFDI e.V.

Baden-Wuerttemberg State Archives (LABW, co-applicant institution): LABW is a leading research infrastructure institution and a centre for regional studies, with decades of experience in digitisation and the online presentation of cultural assets. It is a founding member of the Competence Network of the German Digital Library (DDB), operator of the DDB ingest agency for archives and a member of the DDB Board of Directors. LABW is also significantly involved in the German Archives Portal (Archivportal-D), the main infrastructure for networked research of digital sources and indexing information from German archives. By developing such systems, LABW plays a key role in the conception and distribution of data-exchange formats and standards – EAD (DDB), archival METS/MODS profile – and in the further development of the DFG Viewer. LABW is an active partner for historical research in many third-party funded cooperation projects (e.g. the DFG project “GND4C – GND for Cultural Data”). With DIMAG, LABW has created one of Germany's leading infrastructures for digital archiving. It is constantly developing the system with other federal states and partner institutions. Finally, it is strengthening its activities in the fields of the optimisation of digital research data and RDM through the third-party funded project FDMLab@LABW, which started in 2020.

Bavarian State Library (BSB, co-applicant institution): The BSB is one of the largest European universal libraries and a world-renowned international research library and heritage institution. It forms part of Germany’s virtual national library and is the central state library and repository library of the Free State of Bavaria. Its total holdings amount to around 34 million items, among them almost 10.9 million books, over 54,000 current periodicals, 17 million photographs/images and 140,000 manuscripts. Furthermore, the BSB holds the largest digital data stock of all German libraries with more than 2.5 million digitised works. Around 70 per cent of its copyright-free holdings are freely accessible on the Internet. Besides professional digitisation, the BSB’s own Munich Digitisation Centre (MDZ) develops software and research tools. The MDZ also secures content for future generations through long-term archiving. The BSB’s other sophisticated digital services include four Specialised Information Services and the research data service OstData. The library also operates bavarikon, the internet portal of the Free State of Bavaria. One of the BSB’s departments is the Head Office of the Bavarian Library Network, the regional service and competence centre for all the library network’s members.

Deutsches Museum, Munich (DM, co-applicant institution): Founded in 1903, the Deutsches Museum (DM) is one of the world’s largest museums devoted to technology and the natural sciences. With its main venue in Munich’s city centre and four branch museums, it attracts some 1.5 million visitors a year, with a further 2.5 million visits to its website. Digitisation started early at the DM and developed into the institution's most dynamic and fastest growing research area. The Deutsches Museum Digital, a digital twin of the physical museum, combines extended digitisation of the collections with cutting-edge research in digital humanities. Practical tools such as the VRlab – a Virtual Reality hub to experiment with innovative tools to educate and reflect on

digital storytelling – complement digital research and digital information research infrastructures. DM has been a key player in creating large-scale collaborative digital programs such as the DFG-funded Fachinformationsdienst Geschichtswissenschaft, a specialised information service for Historical Studies run in collaboration with the Bavarian State Library, and KultSam (Cultural Collections as Digital Repositories of Knowledge for Research, Teaching and Public Communication), with the purpose of opening-up the collections for interdisciplinary and transdisciplinary research based on digitisation.

FIZ Karlsruhe – Leibniz Institute for Information Infrastructure (FIZ, co-applicant Institution): FIZ Karlsruhe curates and indexes large amounts of patent information and research data from many sources. Nearly 300 employees develop and operate innovative information services and e-research solutions for precise retrieval and intelligent analysis of such data. FIZ Karlsruhe collaborates with academic and research organisations to conduct applied research and is part of several high-level national and international working groups and committees addressing information infrastructure issues. The department e-Research focuses on RDM, digital long-term archiving and virtual research environments. With a strong background in software engineering, it has participated in or led projects like National Hosting of Electronic Resources, RADAR, German Newspaper Portal, TOPORAZ and Time Machine Europe. It oversees the operation and software development of the German Digital Library (DBB) and German Archives Portal. The research department Information Service Engineering covers semantic technologies, knowledge discovery, ontological engineering and exploratory search. It brings deep experience with the design, implementation and exploitation of ontologies, knowledge graphs and linked data to the consortium. The department Intellectual Property Rights in Distributed Information Infrastructures contributes expertise in copyright, IT (security) and data protection law related to research data on a German and EU level.

The German Historical Association (VHD, Verband der Historiker und Historikerinnen Deutschlands e. V., co-applicant institution) is the largest German scholarly association of professional historians, with 3,400 members. Founded in 1895, the VHD is firmly established in the field and deeply rooted in the research community. Every two years, together with the Association of History Teachers (VGD), it organises one of the largest humanities conferences in Europe, with more than 4,000 participants, providing a vital forum for debate on current research questions, research results and innovative methodological approaches. The VHD’s other main task is promoting the study of history and, above all, representing German historians vis-à-vis the public and political representatives, especially with regard to cultural and academic policy. Since 2004, it has awarded the Hedwig Hintze Prize for outstanding dissertations and the Carl Erdmann Prize for outstanding post-doctoral qualification theses (*Habilitationen*). The VHD covers all historical epochs and sub-disciplines, and currently maintains specialised working groups. Since 2012, it has addressed the digital turn in the humanities through the formation of a working group

on Digital History (VHD-DGW). Since September 2015, the VHD has been instrumental in initiating and developing the 4Memory consortium (see 3.1.2).

Herder Institute for Historical Research on East Central Europe – Institute of the Leibniz-Association (HI, co-applicant institution): Since 1950, HI has combined research on East Central Europe with the mandate of setting up and expanding scholarly infrastructures. It generates many forms of research data and brings together unique collections under one roof: a press-clipping archive (5 million entries), an image archive (750,000 images and born digitals), a map collection (50,000 maps and 10,000 aerial photographs), a musical collection and a further 1,300 metres of historical documents. The collections are complemented by one of the leading libraries on East European history (650,000 media items). By founding a new department for the development of research infrastructures in 2015, HI made a significant step in the direction of digital knowledge management. Since 2021, it has been implementing a research infrastructure “Digital and critical documentation of cultural heritage in Eastern Europe” together with national and international partners. The aim is to map the complexity of historical sources; to further develop existing standard vocabulary; to analyse the interrelations between the growing number of digitised sources and the development of new methods and tools (academic software development) in transnational contexts and to map historicity and source criticism in technical systems in order to show multiple perspectives in metadata structures.

The Humboldt University of Berlin (HU, co-applicant institution): HU is one of the five best German universities according to THE rankings. The digital transformation of science and society and new digital research environments are among its key development priorities. These are pursued through local collaborative projects such as the Weizenbaum Institute for the Networked Society or the Humboldt Institute for Internet and Society (HIIG) and initiatives such as the Humboldt Law Clinic Internet Law (HLCI). With its Computer and Media Service (CMS) and the University Library (UB), one of Germany’s largest, HU offers an advanced infrastructure in Open Science, RDM and Data Literacy. HU projects and offerings include Laudatio (Long-term Access and Usage of Deeply Annotated Information) as a research data repository for texts; FDMentor and FDnext; and the planned Scholarly Maker Space for Digital Humanities in the area of RDM. The Faculty of Arts with its Institute of History (IfG) has a long tradition in Digital Humanities, particularly Digital History. This commitment has been consolidated by the new professorship for Digital History in 2020, which is to be followed by other professorships with a digital focus. This strategy is embedded in the overarching initiative “Digital Humanities Campus Mitte”, which bundles the expertise of HU institutes and faculties in the humanities and social sciences and offers further resources in digital research, RDM and research support.

Martin-Luther-University Halle-Wittenberg, Historical Data Centre of Saxony-Anhalt (MLU-HistData, co-applicant institution): MLU-HistData has been a permanent institution for subject-specific RDM according to FAIR principles since 2008 and regularly conducts consultations on

data management for all stages of the data lifecycle. As an innovative institution for digital transformation, it develops guidelines for subject-specific data management, especially for processes of data curation, authorship, documentation and data quality assurance. HistData advises on research projects and supports researchers as a data publisher through data curation, enrichment with authority data and FDM. MLU-HistData cooperates with the University and State Library of Saxony-Anhalt in expanding the certified data repository shared_it (17,479 historical data sets to date). For protected data, the BOLSA portal was opened in spring 2020 to ensure access to metadata. The data centre is an innovator in the field of entity extraction, modelling of historical data, scientific data analysis and analytical standards data development in cooperation with international standards (PST, HISCO, KldB2010, GND) and international networks (ENCHOS, History for Data, ICARUS). MLU-HistData coordinates tools and workflows for comprehensive data recording together with citizen science communities (Verein für Computergenealogie (CompGen), Networks of national and regional history) and develops large data and standards databases (GOV) with over 20 million individual records in this context. As a member of various competence networks of the DHd Association, GoFAIR, Research Data Alliance and VHD, HistData regularly organises workshops on the methodological development of historical science, data management and events on data literacy. MLU-HistData is a member of BERD@NFDI and NFDI e.V.

Trier University (UT, co-applicant institution): UT has a strong profile in Digital Humanities and Digital History. Its history department is a major national and international player in digital history. Faculty members provide an array of data types, several data-driven research projects, expertise in information technologies and data modelling and a proven record in developing innovative digital research tools, techniques, workflows and infrastructures, such as simulation techniques, network visualisation, topic modelling and time-specific controlled vocabularies (see LoC in the appendix). UT has developed tools and services for collaborative RDM in the humanities. The virtual research environment “FuD” serves the history community in Germany and beyond, offering an infrastructure that combines hermeneutic text analysis, quantitative data analysis and cartographic methods and that can handle and process a range of digital sources. To support humanities researchers developing and implementing IT-based research-projects, UT has created the Service Center eSciences (SeS). It coordinates and implements UT’s RDM strategy and is responsible for RDM-consultations and training. It has developed the DIAMANT-Model, including an RDM competence profile and a qualification concept to implement and support an RDM service landscape in research institutions. SeS has also developed the virtual data repository ViDa for long-term preservation of historical data domains. ViDa will be the long-term archiving infrastructure of Rhineland-Palatinate, whose RDM strategy SeS will co-coordinate. UT is also a participant in the Text+ consortium.

University of Applied Sciences, Darmstadt (h_da, co-applicant Institution): University of Applied Sciences (h_da) is one of the largest universities of applied sciences (HAW) in Germany. It offers over 60 study programmes in arts and sciences for 16,000 students and is one of the few German HAWs with the right to award doctorates in its strong research areas of sustainability sciences, applied computer science and social work. Currently, several third-party funded projects (BMBF, HMWK) are being carried out, addressing the modelling of digital research data in the arts and humanities, the digitisation of cultural heritage, digital scholarly editions and data literacy. This research is addressed primarily from an information and library-science perspective, and h_da is active in several national and international professional associations such as KIBA and IFLA. Beyond 4Memory, h_da is involved in several national research infrastructures, e.g. in the humanities-oriented consortium CLARIAH-DE, and it is a participant in NFDI4Culture, Text+, and NFDI4Ing. Building up research data infrastructures is central to its interests, as h_da, like all other Hessian universities and HAWs, is involved in the state-wide research data initiative HeFDI. In this framework, alongside developing digital RDM services, further education and training courses for data literacy are being established. This field is located at the transdisciplinary interface between engineering and the humanities on the one hand and information-, computer- and library science on the other.

3.1.2 Embeddedness of the consortium in the community of interest

4Memory is firmly embedded in the historical disciplines and institutions. Its community includes individual scholars, projects, institutions, professional societies, networks and publishers and extends beyond academia to include users of historical data in the context of citizen science.

4Memory: a product of its community

4Memory has developed over four years through an intensive, continuous dialogue with its community. The German Historical Association (VHD), a co-applicant, has played a crucial role. In its Working Group on Digital History (VHD-DGW) and regular events, the VHD has made addressing the opportunities and challenges of digitalisation a priority. In 2017, the VHD issued a position paper on the creation of the NFDI, setting out its support and the necessity of tailoring it to the needs of historical research (VHD 2017).[2] The VHD has since co-sponsored key organisational meetings, such as a conference on “Research Data in History” and the workshop “History in the NFDI Process” in 2018. It has intensively informed its members (via its mailing lists, journal and conference events) about the NFDI and 4Memory. The central professional society for German historians will remain a key actor in 4Memory, keeping it community driven: the VHD has enlarged its board by a delegated member of our consortium. Other associations have joined

[2] Discussions with the VHD also led to a position paper from fellow co-applicant BSB (BSB 2018).

our initiative, representing the broad community of the historically oriented humanities as participants (GSWG, MG) or agreeing to collaborate as partners.[3]

4Memory has grown through dialogue and maintained the greatest degree of transparency. In 2019 the applicant institution (IEG) issued a call for participation and an accompanying survey of RDM needs and possible contributions. The call was widely promoted and crucially defined our work programme. With the launch of the 4Memory website (https://4memory.de) in late 2019, a presence on Twitter and an electronic newsletter, the initiative entered into a more direct contact with our community, a process that has continued in our "community workshops" since early 2020, accompanied by publications in journals from the community.[4] The global pandemic of early 2020 led to a planned series of events throughout Germany in March and April being transformed into virtual meetings. More than 150 participants took part in two days of online discussions at the end of March that were moderated by our task-area coordinators. The edited questions and answers were placed online in a downloadable 40-page document, making the results available to the community.[5] We issued an open letter explaining our plans and the latest developments in our formation.[6] A follow-up workshop and a call for "problem stories" – short statements of the everyday challenges and needs faced by those working with historical research data – clarified issues relevant to community members. Nearly 100 submitted problem stories were systematically analysed and taken into account in our work programme. We published a preliminary outline of our planned structure on our website: this provided an important aid for further intensive discussions with participants, impacting the consortium's overall orientation. Our discussions crossed disciplinary boundaries: consultation with representatives of various area studies led us to critically reflect on taking into account the issues and perspectives of non-European providers and users of research data. In preparing our revised application we have continued our open dialogue with our community through further consortial meetings, social media and direct discussions at the task-area level.

4Memory: Continuous community engagement

We will remain embedded in our community through effective communications built into 4Memory, from individual tasks up to our governance structure. Our TAs have integrated measures to maintain contact with our community, e.g., surveying and assessing data-quality needs (TA1-M1), pursuing collaborative development of data and documentation standards (TA2), setting up counselling services on interoperability and reusability (TA3-M4) and digital research literacy

[3] The partnerships already finalised include the Digital Humanities (DHd), Philosophy (DGPhil), Art History (VDK), Religious Studies (DVRW), Information Studies (HVIW), Egyptology (IAE), Ethnology (DGV), East European Studies (DGO, VOH), Research Libraries (DBV, LIBER, CERL).
[4] The articles discuss the role of archives (Maier, Fähle, and Neuburger 2020), new knowledge orders and data quality (Schlotheuber and Paulmann 2020), our LINKAGE goals (Paulmann, Wood, and Cremer 2020), and integrating research, memory and infrastructure institutions (Paulmann and Wood 2020).
[5] https://4memory.de/fragen-und-antworten-zu-4memory/
[6] https://4memory.de/2020/04/23/open-letter-to-the-4memory-community/

(TA4-M4) together with community representatives. TA5 is devoted to encouraging a new data culture in the historically oriented humanities, with tasks centred on community engagement, such as promoting a discourse on the benefits of using and reusing research data and on digital methods (TA5-M2). Our work programme includes communications with institutional actors, such as Specialised Information Services,[7] collaborative research projects,[8] federal-state RDM initiatives, memorial sites and citizen science organisations, including the umbrella organisation of over 200 German historical societies (GdGA).[9] TA6 ("Participation and Steering") enables stakeholder involvement in our continuous development. Governance procedures will give the community a voice in the consortium (see 3.4): our Community Forum will reflect community diversity in terms of organisational types (individual researchers, professional societies, representatives of memory institutions, etc.) and the disciplines of the historically oriented humanities; it elects representatives to the General Assembly. The community will be represented in the Academic Advisory Board and continually involved in our internal evaluation procedures. Our community is also connected to us through partnerships – cemented through letters of support – that will continue as we develop.[10]

3.2 The consortium within the NFDI

4Memory is the only planned consortium defined by its focus on the distinct challenges of historical research and historicity of data: our expertise in hermeneutic approaches accounting for multiperspectivity and the challenges of fragile, incomplete data will contribute not only to community-specific requirements but also point to numerous interfaces with other consortia. A majority of our co-applicants and many of our participants have become institutional members of NFDI e.V., providing an additional avenue for the consortium's contributions.

4Memory has a strong, complementary connection with three other humanities consortia, each of which addresses a specific community: **NFDI4Culture**, **NFDI4Objects** and **Text+**. In 2019 we signed a joint Memorandum of Understanding (MoU) in which we pledged to collaborate on issues of common interest. In 2020 we signed a revised version of the MoU (Brünger-Weilandt et al. 2020) that deepened our plans for collaboration. The memorandum group meets regularly to discuss how services and tools relevant to the humanities as a whole can be jointly developed

[7] DFG Specialised Information Services (SIS or FID) play key roles in the NFDI (Harbeck and Kaun 2019; AG FID 2018). Participating or cooperating FIDs in 4Memory include "African Studies", "Classics (Propylaeum)", "CrossAsia", "Educational Science and Educational Research", "History", "Jewish Studies", "Latin America", "Mid-East, North-Africa and Islam Studies", "Social and Cultural Anthropology" and "Russian, East and Southeast European Studies". See Appendix 1.2 for references.

[8] Collaborative research projects include the DFG programmes Collaborative Research Centres (e.g. SFB 1199), Clusters of Excellence (e.g. EXC 2176), Long-Term Projects in the Humanities and Social Sciences; the BMBF-funded Merian-Tagore International Centre (New Delhi) and the Marbach Weimar Wolfenbüttel Research Association; and projects of the Academies of Sciences and Humanities (e.g. Germania Sacra).

[9] The historical societies ("Geschichtsvereine") form the largest group of more than 130,000 historically engaged citizens with regional or locally specialised expertise, autobiographical knowledge, special language and writing skills and dedication to historical sources. This is also true for the community of our other key citizen science partner, the Wikimedia Foundation.

[10] https://4memory.de/partners-2/

and operated and to coordinate our joint contributions to the four prioritised cross-cutting sections named below. 4Memory has, furthermore, signed the "Leipzig-Berlin-Erklärung" (Maik Bierwirth et al. 2020) on cross-cutting topics. On a bilateral basis we have agreed to collaborate particularly closely with **NFDI4Culture** on copyright in historical sources, knowledge graph implementations, integrating data and code literacy in curricula, data quality management and editorial processes for data publication; with **NFDI4Objects** on provenance and research ethics, interoperability through authority data and ontologies, services for discovery and long-term storage, data epistemologies; and **Text+** on data ethics in historical data, integrating authority data, data literacy in text-related tools, data quality assurance and interoperability in metadata infrastructures. In these collaborations our complementary areas of expertise – which, taken together, represent a wide spectrum of the humanities – will be of great value to the NFDI.

Beyond the named humanities consortia, 4Memory will cooperate with other consortia regarding knowledge exchange and cross-cutting topics. 4Memory will collaborate with **BERD@NFDI** on topics such as automated text, layout and structure recognition as well as in the harmonisation of terminologies and the mutual exchange of controlled vocabularies and ontologies. Historical data have many of the same issues as the qualitative and quantitative social science and economic data that are the focus of **KonsortSWD**. **MaRDI** supports the network analysis of historical data (such as correspondence or itineraries) and develops algorithms that take into account temporal issues and factual inaccuracies, capabilities that are of interest to 4Memory. We will collaborate with **NFDI4Ing**, particularly with its task area "Base Services", within "Quality assurance in RDM processes and metrics for FAIR data" on advancing Research Data Management Organizer (RDMO) and its discipline-specific usage within our community (see also 4.1.3). We also see potential interfaces with basic services with regard to identity and access management and long-term archiving, metadata handling and workflow automation.

The distinct objectives and experience 4Memory brings together will enable us to make unique contributions to the following cross-cutting sections:

- (Meta)Data, Findability; Terminologies; Provenance
- Common Infrastructure
- Rights and Ethics
- Training and Education

Networking with other consortia has been facilitated by the fact that many members of 4Memory are also active in other consortia. The following NFDI consortia include members from 4Memory: **BERD@NFDI**: GSWG, LRZ, MLU-HistData, UBM, WA-VfS, ZBW; **DataPLANT**: ZDV-T; **FAIRagri**: FIZ; **FAIRmat**: HU, LRZ; **GHGA**: LRZ; **InnoMatSafety**: FIZ; **KonsortSWD**: DIPF, UBM, UL-UB, ZBW; **MaRDI**: FIZ; **NFDI-MatWerk**: FIZ; **NFDI4BioDiversität**: GU-UB; **NFDI4Chem**: FIZ; **NFDI4Culture**: BAdW, BSB, DDK, digiCULT, DM, DNB, FIZ, FSU-II, GU-UB, h_da, HI, SBB-PK, SLUB, SUB, UL-UB; **NFDI4DataScien ce**: FIZ, ZBW; **NFDI4Earth**: LRZ, ZDV-

T; **NFDI4Ing**: h_da, LRZ, SLUB, THK; **NFDI4Microbiota**: THK; **NFDI4Objects**: BAdW, DNB, SBB-PK, SUB; **NFDI4Phys**: FIZ; **NFDIxCS**: FIZ; **PUNCH4NFDI**: LRZ; **Text+**: BAdW, BBF, DNB, GU-UB, h_da, HAB, OFU, SBB-PK, SUB, UHH-FDM, UT, the German Humanities Institutes Abroad via the Max Weber Foundation.

3.3 International networking

Thematically and institutionally, historical research possesses a “natural” scope beyond the nation. 4Memory’s community and its participants and co-applicants have extensive, longstanding links with academic communities in other countries through German institutions within Germany devoted to understanding other nations and regions and German institutions based in other countries and integrated into their local research cultures. They are platforms for cooperation and catalysts for intercultural and transcultural knowledge production. They have all built up digital resources and advanced digital research in their respective domains, developing experience and defining further needs that 4Memory has integrated into its work programme. Within the first category, 4Memory’s co-applicants and participants include research institutions and libraries (e.g. the Leibniz Institute of European History, the Herder Institute for Historical Research on East Central Europe, the Ibero-American Institute, the Centre for Asian and Transcultural Studies, the Leipzig Research Center Global Dynamics, the Global and European Studies Institute); library services (e.g. Specialised Information Services on, among others, African studies, Asia, Jewish studies, Middle East-, North Africa- and Islamic Studies, Northern Europe, Russian, East and Southeast European Studies, and social and cultural anthropology); and collaborative interdisciplinary research projects (e.g. the Cluster of Excellence Africa Multiple). Within the second category, our participants include the German (Historical) Institutes in Beirut, London, Moscow, New Delhi, Paris, Rome, Tokyo and Washington.

4Memory and its co-applicants are involved in several international networks developing digital methods, data management and data literacy in historical research. This includes RDM initiatives (Research Data Alliance, GO FAIR); networks of memory institutions (LIBER, CERL, ICOM, ICARUS); standardisation committees (CIDOC, Resource Description and Access, MODS/MADS, IIIF); ESFRI consortia (DARIAH-EU, EHRI, OPERAS, and RESILIENCE, including contributions within the SSHOC/EOSC activities; digital history networks (Data for History, ENCHOS). Furthermore, 4Memory seeks close cooperation with the renowned digital history centres C2DH (Luxembourg), Huygens ING (Amsterdam), and RRCHNM (Washington DC), which offer 4Memory international perspectives and pathways for knowledge exchange.

Two examples can illustrate the benefits accruing to our consortium (and thus to the NFDI) from such forms of collaboration in the digital humanities: the field of **Russian, East and Southeast European** history and that of **area and global studies**. The first example of Russian and East European history demonstrates how 4Memory combines national, transnational and global

aspects of **collaboration on research data and data literacy**. Substantial transnational cooperation on **data curation and creating knowledge spaces** exists in areas such as **legal and ethical standards**, **multilinguality and multiperspectivity** and **decolonising Western norms and narratives** in research data on Eastern Europe. The large number of German-language historical sources also requires cross-border cooperation on **thesauri, metadata management, data literacy and culture-sensitive methods of source criticism.** The complex entanglements of nations and communities as well as Germany's violent and expansionist pre-1945 policies require a context-sensitive handling of research data and advanced metadata strategies based on transnational cooperation. Two of 4Memory's co-applicants will coordinate a network to produce deliverables and recommendations for data management in this field: the BSB and the HI (with the support of DHI Moskau as participant and DHI Warsaw as partner). They bring together collections, digital resources, firsthand expertise and collaborations with key institutions in their respective countries, the most important here being the Collegium Carolinum in Munich, the Leibniz-Institute for East- and Southeast European Research in Regensburg, the Leibniz Institute for History and Culture in Easern Europe (GWZO) in Leipzig (already cooperating in the research data service OstData) and the Research Centre for East European Studies at the University of Bremen.

For our second example, substantial parts of the fields of area and global studies belong to the historically oriented humanities, sharing a common set of methods and interests but also offering expanded perspectives on research data beyond Europe. As expressed in our LINKAGE aims, we are committed to critically adapting the humanities' present knowledge order by accounting for non-European concepts. Prominent among these issues are questions of **data ethics and rights arising from the colonial and postcolonial past**, varying **traditions of knowledge production** and **power differentials in the present world order**. Many practical and technical issues arise regarding handling **non-European languages in digitisation processes**, **access to archives** and **varying degrees of digitalisation** in different regions' academic systems. 4Memory has from the beginning sought out perspectives on data and data management beyond those German-speaking and European actors who necessarily make up the vast majority of our participants and partners. Through discussions with representatives of fields such as Area Studies (in Latin America, Africa, the Middle East, South Asia and Asia), Specialised Information Services (FIDs) in the globally oriented topic areas mentioned above and internationally based German institutions, we have imbued our efforts with a **critical (and, where necessary, self-critical) perspective on historically oriented research data**. These connections will continue across the first five-year phase, part of an ongoing process that will start with intensive exchanges among a smaller group of participants but expand on the firm foundations that are thereby established, e.g. discussions with region-specific FIDs to incorporate the perspectives of various world regions. We are well placed not only to improve the quality and international relevance of

our consortium's efforts but also to ensure that we can **take into account the CARE principles** (GIDA 2019).

3.4 Organisational structure and viability

NFDI4Memory implements a **process-oriented governance structure** to serve its members and community. The architecture of participation and steering is guided by two principal missions: 1) enabling participation, advice and evaluation from our community and 2) managing the connections within, the services for and steerage of our consortium. Accomplishing these missions requires meeting key aims. First, the governance structure **enables exchange** between our community and the consortium as well as internal communication among the co-spokespersons, participants and the consortium's various bodies. Built upon a needs-based approach, the governance process creates opportunities for mutual feedback and avoids redundancies. Second, the structure ensures the **transparency and efficiency** of decision making, a crucial aim for a consortium serving a large community with manifold interests and varying resources. Third, we create a process-oriented structure with **multiple evaluation procedures** on levels extending from the community as a whole to our executive agents: this structure will **allow services to evolve and innovative tasks to emerge** from its own work and from the integration of new requirements and agents in the first and second phases. The responsibility for the governance process rests with TA6 (Participation and Steering).

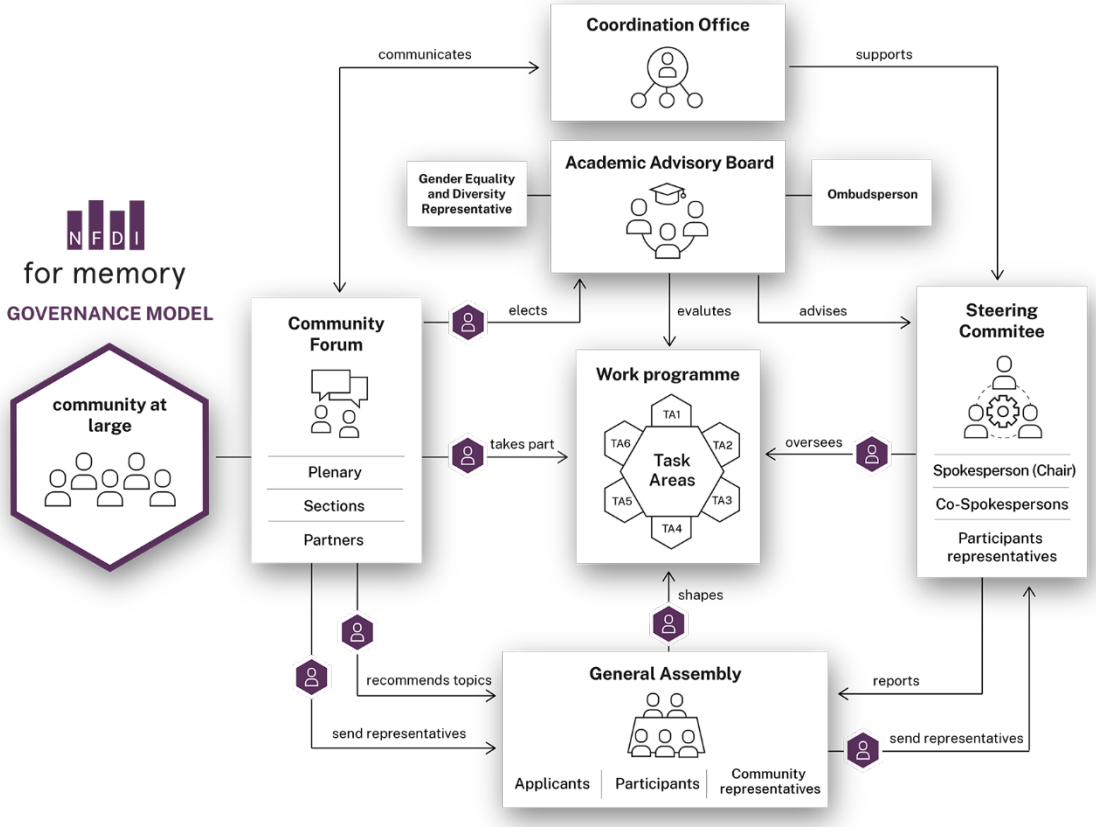

Figure 2: 4Memory governance model

Our governance enables significant community involvement, giving a formal voice to individual scholars, those working in large research projects and institutions ranging from universities and memory institutions to learned societies and infrastructure institutions. Community members may engage with the consortium in many ways: as participants, as partners or more informally through public forums and events. The governance process rests on the activities of consortium members; its structure also takes account of the NFDI as a whole and other consortia within it.

The governance process has four domains: (A) community involvement; (B) representative bodies of the consortium; (C) executive agencies; (D) mediation and conflict resolution.

(A) **Community involvement** in 4Memory ranges from continual collaboration **in its tasks and work programme** to periodic participation in **community events.** The community is formally integrated on several levels of our **representative and executive bodies** through elected representatives, the reception and discussion of annual reports and the evaluation procedures.

(B) Three **representative bodies** form the core of the governance process: the 4Memory Community Forum, and the General Assembly and the Academic Advisory Board.

(1) The 4Memory Community Forum is open to the community and meets annually. It includes individual researchers and partners from learned societies, collaborative projects and institutions. Possibilities for virtual participation will be taken into account. The Community Forum is organised in plenary and thematic sections. Forum sections discuss topical or methodological issues; they also discuss and evaluate the measures undertaken by our TAs and make recommendations to the General Assembly regarding needs and improvements. The Forum's plenary receives and discusses the annual report by the Spokesperson and the 4Memory Steering Committee. 4Memory partners elect community representatives to and propose agenda items for the General Assembly. Partnerships are established by formal letters of support and approval by the General Assembly.

(2) The General Assembly comprises all the consortium's members (co-applicant and participant institutions) and community representatives elected by the partners in the Community Forum. The General Assembly receives, discusses and approves the annual report by the Spokesperson and the Steering Committee. It discusses the strategic planning of the consortium and its task areas based on recommendations from the Community Forum, the TAs' achievements and the advice of the Academic Advisory Board. It approves partnerships of 4Memory. It may form temporary and permanent working groups on strategic objectives, such as innovation, internationalisation, data ethics or technical standards.

(3) The Academic Advisory Board represents the community, learned societies, participants and international partners. It functions continuously and meets annually. It is elected by the Community Forum. To ensure immediate viability, half its ca. twelve members will initially be appointed by the Steering Committee before the consortium's expected starting date. The

Board has advisory and evaluative functions. It advises the Spokesperson and the Steering Committee on our innovation strategy, community involvement, ethical issues, internationalisation and 4Memory's relationship to the NFDI as a whole. It will conduct an audit during the first phase and may evaluate and make recommendations to particular TAs as required. It will select the Ombudsperson from among its members.

(C) The **executive agencies** of NFDI4Memory are the Spokesperson, the Steering Committee, and the 4Memory Coordination Office (4MCO).

The **Spokesperson** represents the consortium externally (within the NFDI, to our community, to the general public) and chairs the Steering Committee. He or she takes responsibility for the consortium vis-à-vis the DFG and coordinates TA6. Five co-spokespersons (one each from Task Areas 1–5) act as deputy spokespersons on a rotating basis.

The **Steering Committee** is the executive agency overseeing the implementation of our work programme. It has the responsibility for the realisation and strategic evolution of the measures; it oversees, evaluates and decides upon changes to the consortium's strategic development. The Steering Committee approves the disbursement of funds to participants. Following proposals from the task areas, and in consultation with the Academic Advisory Board, it makes decisions about the use of the centrally administered strategic fund (TA6-M3-T1) in accordance with 4Memory's objectives and the evolution of its work programme. The Steering Committee consists of the Spokesperson, the Co-spokespersons from the applicant institutions, representatives of the participants, and the consortium's Managing Director. Among the operational-level executive bodies are the officers of the **4MCO.** Their roles, such as with regard to the conclusion of the initial cooperation agreements with the members of the consortium and the evaluation and quality assurance, are described in detail in TA6's work programme, 5.6.1.

(D) Processes of mediation and conflict resolution are handled by an **ombudsperson** who will have the roles of addressing complaints and mediating internal disputes. The ombudsperson is part of the Academic Advisory Board and chosen by it from among its members. From its membership, the Board will also select a person responsible for **gender equality and diversity management** tasked with ensuring these issues are consistently addressed in the consortium.

3.5 Operating model

NFDI4Memory will integrate into the NFDI governance structure, which is being established at the time of writing. It will not transform into a legal entity of its own. Its operating model will be based on a consortium contract, as is the practice in national and international consortium projects. The consortium contract will be in accordance with the template provided by DFG. Following the template's guidelines, our activities will not be based upon the commercial exchange of goods and services. Funds will be transferred only between non-profit entities or legal entities under public law, i.e. the members of the consortium, and only by way of a genuine

grant. They will serve the sole purpose of promoting scholarship and research. A majority of 4Memory's co-applicant institutions and many of its participants have become members of NFDI e.V. and are already actively shaping the organisation's work (through, for example, contributions to cross-cutting sections).

4 Research Data Management Strategy

Our revised RDM strategy (a version of the original was published in open access as Cremer et al. 2021) is the fruit of ongoing improvements based on feedback from the colloquium, external reviewers, expert committee, consortium members and our community. Beyond strengthening all sections, we have more clearly formulated our vision of innovative digital history, our strategy for metadata harmonisation and our critical approach to research ethics, and we have increased our focus on competence building and concrete actions in engaging with the NFDI.

4.1 State of the art and needs analysis

4.1.1 Research data in the historically oriented humanities

Historians have always been data collectors, creators, curators and reusers. Their data consist of the records they seek, gather, validate, annotate, compare and represent in contextualised formats to answer research questions and to support statements, methods, arguments, hypotheses or theories. Research-driven data aggregation thus creates data collections reflecting specific interests and methods. These collections contain data of various provenance – from memory institutions, state agencies, private persons, companies, citizen-science organisations[11] and others – and comprise a range of document types and of data and file formats, many of which have yet to be digitised professionally or in standardised ways.

Historians have yet to reach a consensus on defining "research data". They have traditionally distinguished between the sources from which they extract historical knowledge and the information they create through research guided by particular questions, methodological assumptions, processes and perspectives. The growth of digital RDM has caused many historical researchers to describe the historical record itself – i.e. the sources – as "research data" (see surveys in Germany: Andorfer 2015, 13–14; Queckbörner 2019, 31). The definition of historical sources as "research data" is not without controversy, but it has been tentatively established that digital representations of historical records are fitting subjects for RDM. In essence, "data" now encompasses all forms of representation, from the analogue object to the reused data produced

[11] Examples from the consortium are the collection of 185,000 maps of Perthes at Gotha (memory institution), the German documents of the First World War in the Central Archive of the Ministry of Defense of the Russian Federation at DHI Moskau (governance institution), the 15 million photographs from the german weekly news magazine *Stern* at BSB (private property), the 1.2 million entries on historical places, districts, towns, counties, regions, etc. in GOV from CompGen (citizen science).

during research. In this process historians traditionally create source excerpts, transcriptions, bibliographical and biographical records, finding aids, inventories, interviews, statistical data (e.g. time series), and/or structured information on events, facts, places or persons, including the provenance and context of any of this information.

The RfII calls research data “any data that is generated in the course of scientific work, for example through observations, experiments, simulations, surveys, questioning, source analysis, records, digitalisation, evaluations”, including “such data which are not acquired by science itself, but which science accesses for research purposes in order to use them as a methodologically necessary basis for the concrete research process” (RfII 2020c, 106). Adapted to history, this adds new data types to the categories of historical sources and published results, broadening the definition of research data, as the VHD’s position paper on the NFDI elaborates (VHD 2017, 3–4). This wider understanding leads to a **general definition of research data in the historically oriented disciplines**: any information, regardless of its provenance, is considered “research data” as soon as it is collected, described, analysed, evaluated and/or created and stored in machine-readable form for the purpose of maintaining the traceability of research results or for archiving, citation and further processing. Hence, research data include all media formats and digital representations of analogue sources – publications, sources, metadata, controlled vocabularies, statistical or simulation data, etc. – as well as any information that can be used to answer research questions (Oltersdorf and Schmunk 2016).

The issues named above are common to the humanities, which share generic RDM needs (Schöch 2013; Puhl et al. 2015; Cremer, Klaffki, and Steyer 2018). Research data in the humanities are gathered and captured – in a process guided by specific aims – as an act of creation and construction (Drucker 2011). Sources are potentially endless, since nearly any record of human activity or experience is a potential data source (Borgman 2007, 216). However, **historical data pose specific challenges** to the historical method in digital scholarship, which must be addressed collaboratively in the ***N****etwork of historically oriented research communities*. First, the historical record’s incompleteness and separation from its original context makes establishing sources’ authenticity challenging. Digitisation often adds another layer of difficulty, as digital representations lack the physical attributes normally used to prove – or question – authenticity. Second, data collections combine layers of complexity. Sources vary in format and structure, with multimedia collections and unstructured items as the norm: their human-created contents are ambiguous, even contradictory. Third, establishing data as a legitimate research output requires contextualising information regarding its historical origin, provenance through data processing and form of representation. Larger collections and digitised formats, however, have encouraged decontextualisation. Finally, historically analysing sources requires extending source criticism – from production and provenance to the different representations of data – reflecting the “dual nature of digital data as both content *and* form” (Barats, Schafer, and Fickers 2020).

In particular, born-digital sources pose yet unsolved problems, since the digital conditions in which they are produced mean traditional methods relevant to analogue materials cannot be applied. Born-digital sources are fundamentally modifiable and manipulable. Bound to specific digital frameworks and platforms, they are often volatile and ephemeral – not only in their presentation but also in their content and substance – and subject to continuous change through digital processes. Due to hypertextuality and technical dependencies, they can often no longer be described and processed as independent objects detached from their environment (Föhr 2017, Fickers 2020). New forms of authorship have to be taken into account, such as collaborative writing or algorithmic authorship (bots). New practices of reusing and modifying or decontextualising and recontextualising content (as a whole or in part) pose particular challenges: through copying, editing, mashups, memes and deep-fakes the original content ultimately dissolves (Stalder 2018). There is also a massive amount of digital-born data of utmost importance for future historical research (e.g. web pages, SMS, chat histories and social media) that either ends up lost or deleted or that can only be accessed – let alone used – under legally restrictive conditions (Brügger and Schröder 2017, Milligan 2019). The development of historical methods and their critical reflection will shape the humanities' ***K****nowledge order for the digital future of the past*.

4.1.2 Scholarly principles in the historically oriented humanities

Historical research, regardless of its degree of digital transformation, is defined by its techniques. The historically oriented humanities share **scholarly principles** with the humanities overall (Unsworth 1997), but domain-specific aspects are summarised in the white paper "Digital History and Argument" (Arguing with Digital History working group 2017):

- *Selection*: In addition to the processes of capturing and creating data, researchers identify the relevant sources and select them as their data. This selection is based on "the truthfulness of a source, its aesthetic qualities, its representativeness, or its uniqueness."
- *Synthesis*: To validate statements, methods, arguments, hypotheses or theories, selected sources and further created data are analysed for patterns and structures through modes of composition and comparison, arranged according to time, place, topic or other principles.
- *Contextualisation*: Creating an accurate, reliable and persuasive account of the past requires a complex picture of an event, including its environment. The scale of this context varies: from biographical to global history or from microhistory to perspectives spanning centuries.
- *Communication*: Representing the results conveys the information about the past and the argument and describes the processes of selection, synthesis and contextualisation.

A full approach to RDM supports these principles, building them into workflows and the data lifecycle (all TAs), enhancing selection possibilities by extending access to sources and detailed descriptions of them (TA1, TA2, TA3); and realising the potential of public interaction (TA5).

According to the ENUMERATE 4 Survey (Nauta, van den Heuvel, and Teunisse 2017), about 50 per cent of analogue heritage collections still need to be digitally reproduced, while only about 10 per cent are already digitised. Much historical source material will remain **analogue research data**. As Zaagsma notes, “The real challenge is to be consciously hybrid and to integrate ‘traditional’ and ‘digital’ approaches in a new practice of doing history” (Zaagsma 2013, 17). There is a danger, though, “that our narratives of history and identity might thin out to become based on only the most visible sources, places and narratives” (Edmond 2016, 11). Our RDM measures must thus seek to ***A**dvance the analogue / digital interface of historical source material* and reflect hybrid approaches in mixed research data collections and combined methods (all TAs). Linkage between digital data and analogue sources (TA2, TA3) has to be improved. The digitisation of heritage must be promoted through a systematic-programmatic strategy by memory institutions and a research-driven effort through project funding (TA5, TA6).

4.1.3 Research workflows in the historically oriented humanities

Historical research ranges from individual research and groups in smaller projects to large, collaborative and interdisciplinary projects and from using traditional methods on analogue sources to using computational methods on digitised and born-digital data. In the **traditional workflow**, researchers collect sources (often only available in analogue form) in archives and libraries; analyse the selected corpus and answer research questions through established methods. The resulting publication, usually a printed book or peer-reviewed journal article, explains the selection criteria, interprets the sources and cites their (physical) location and includes a source list describing the collections used. While this standardised approach guarantees transparency and intellectual comprehensibility, it limits the reuse of data to the interpretations and references made therein; also, the documentation, contextualisation and enrichment of sources is not transferred back to the memory institutions holding the original sources. Since the traditional project is still the norm in the humanities, 4Memory complements its innovative impulses with RDM measures to enrich analogue workflows with digital services and tools (TA2 and TA3); advance community practices toward standardised RDM (TA1 and TA4); and develop channels and opportunities to integrate essential concepts and practices into the community (TA5).

A growing number of projects over the last decade have implemented **digitised or digitally transformed workflows**. Researchers work together with memory institutions to digitise original sources, enrich them with basic metadata, make them available as data and guarantee long-term preservation of digitised material. At times, an institutional RDM service centre assists the project, serving as stewards for RDM, as described in the project’s data management plan. Challenges remain for traditionally trained and skilled researchers (Hohls, Prinz, and Schlotheuber 2016), as collaborations among researchers, memory institutions and infrastructure partners are mostly insufficient (Lemaire et al. 2020, 6–7). To address such challenges collaboratively, ***L**inking*

research, memory institutions and infrastructures is central to all TAs' activities. Our RDM measures promote digital workflows that enhance research data quality and sustainability (TA1, TA2, TA4, TA5); foster RDM knowledge in historical sciences as a core competence (TA4, TA5); and increase the interoperability of the tools and information systems used in order to create a consistent framework within the research domain (TA2, TA3).

While improving traditional research, "**digital history**" enhances and expands historical methodologies by integrating computational methods, e.g. natural language processing, computer vision, text mining, network analysis and the fields of data science and knowledge engineering (e.g. Romein et al. 2020, Blaney 2021). These methods also add simulations, visualisations and research software to historical data types. In future, historians should be able to work digitally at every stage of the academic research process. In the heuristic phase they should be able to find and secure relevant sources and data on the basis of aggregated inventories, normative vocabularies and knowledge graphs (and, if necessary, be able to commission their digitisation and convert them into text). In the phases of evaluating and enriching the sources and that of analysing them with data-driven methods, researchers should be able to integrate further data from a common data space (for contextualisation, comparison, generating training data but also for machine-learning models, etc.) and be able to rely on a common knowledge graph (e.g. for named entity linkage). Finally, in the phase of securing the results they should be able to publish the findings in digitally enhanced or computational publications and present their data under appropriate quality assurance in suitable repositories in order to make them available as FAIR research data for further reuses.

A transitional term, "digital history" treats born-digital data as a new type of source, developing digital methods for analysis and new forms of academic publishing (Zaagsma 2013); it is both a scholarship of *transfer* (developing and evaluating methods from other disciplines) and of *transformation* (through which these methods remake historical study) (Rehbein 2018, 31). In addition, the mode of work in digital history is shifting towards concepts emphasizing the collaborative nature, such als DH Labs, which also need structures enabling the interdisciplinary exchange of practices (Fickers and van der Heijden 2020; Kemman 2021). As digital methods evolve and become more common, the meaning of "digital history" will become more diffuse, and historians will more precisely define their practices and vocabulary (Crymble 2021). Using digital data, tools and methods brings dependencies on information technology that must be incorporated into historical knowledge production as "digital hermeneutics" (Koolen, van Gorp, and van Ossenbruggen 2019; Fickers 2020; Romele, Severo, and Furia 2020) and their epistemological consequences must be discussed (Hiltmann et al. 2021). RDM measures require a continuous dialogue in the community (Ma and Xiao 2020) on the application of digital methods (all TAs); on bringing data literacy into historical methodology to further ***I****ntegrate historical source*

criticism into data services (TA4); and on digital knowledge production, including ethical and legal implications from a global perspective (TA5, TA6).

In all historical workflows described above a **data management plan** (DMPs) will be critical to implement good scientific practice on research data. Funding bodies have increasingly advocated their use (Smale et al. 2018), and they may soon become mandatory elements of research proposals also in the humanities, even if they are still rarely used outside of larger collaborative project proposals. 4Memory considers DMPs to be a key RDM instrument, and their formal design has been elaborated and adapted to the humanities (Minn and Lemaire 2017). Thus, our RDM measures will use generic approaches (RDMO), also within the NFDI (TA6); will focus on the domain-specific content of DMPs through quality (TA1) and standards (TA2); and will integrate DMPs into research workflows (TA4) and data culture (TA5).

4.2 Metadata standards

4.2.1 Metadata and contextual indexing in the historically oriented humanities

Metadata's relevance for historical research is best expressed by Drucker, who notes that few other kinds of information have a "greater impact on the way we read, receive, search, access, use, and engage with the primary materials of humanities studies than the metadata structures that organise and present that knowledge in digital form" (Drucker 2009, 9). Metadata describing analogue collections and sources – the catalogue of bibliographic records or inventory items, the oldest standard of archives, libraries and museums to make documents findable and accessible – has long been essential to the historical method as a way of referencing and subject-indexing sources. Researchers add significant (meta) information to their data by describing sources or objects or by annotating digital representations on an item level that were previously indexed on a collection level (or as an archival box). In addition to noting an item's attributes, an index of the content is created by assigning headwords, subject terms and categories; by identifying entities; or by providing abstracts. The scholarly principle of *selection* draws relationships to other items within a collection and beyond it; the principle of *contextualisation* enables positioning items within topic-oriented, temporal or spatial scales.

The traditional work by memory institutions in modelling metadata and the enrichment of data by researchers have been **two ways of creating context data that are now supplemented by innovative digital methods for automatic generation of metadata and authority data**.

4.2.2 Concepts for interoperability and harmonisation

As noted, information on historical sources is mainly created to support arguments and provide the basis for high quality metadata useful for discovery and usability. Especially in the digital space this enables serendipity effects during search and comparative and computational approaches during analysis. 4Memory pursues free access to and reuse of research data: rich

metadata serving intellectual reuse and improved retrievability is crucial. We see ourselves as a community of practice strengthening approaches to metadata that address research-specific needs and improve community-specific use. The following approaches are especially important.

Historical bibliography

Specialised or subject bibliographies are still central in the historically oriented humanities for referencing and researching literature. In digital form, they are no longer closed but can potentially become open systems. Digital bibliographies often contain not only bibliographic records but also offer direct access to research literature, and the bibliographically enhanced and high-quality metadata are also available as open data to other reference systems, such as union catalogs (Kraus and Stöhr 2015; Herr, Hofmann, and Getschmann 2018). 4Memory thus aims to contribute to cataloguing rules for research data within its international networks and include research data as well as the respective documentation as independent publications in relevant specialised bibliographies, linking them to published research syntheses (i.e. monographs and articles). This goal requires that research data sets be appraised as a qualified research output. Combining traditional publications and research data in bibliographies will increase the recognition of data publications as respected scholarly achievements. Research syntheses will gain added value by being linked to the data on which they are based. Our transferable approach holds significant potential for other humanities disciplines.

Authority data, vocabularies and taxonomies

As a part of research-guided entity recognition, authority data is particularly relevant for describing objects with metadata and demonstrating the dissolving boundaries between data and metadata: despite the widespread adoption of uniform principles of spelling and writing in the last century, the sources used by historical disciplines contain named entities rich in variants and dialects and, of course, extend across countless languages. Preserving these original variations is vital to historical understanding and source criticism; however, their heterogeneity poses challenges to making them machine-readable. To be useful as research data, therefore, such variants must be provided with ideally persistent identifiers to enable entity referencing and allow links through ontologies, multilingualism and Linked (Open) Data (Estermann et al. 2020). Thus far, the GND, as a standard for German-language authority data with connections to international authority data systems, mainly offers generalised concepts that are often too broadly defined for scholarly applications in historical fields (Moeller 2019, 26–29). This is part of a wider issue: while researchers need specific classifications or categories, memory institutions' indexing tools rely upon generic forms of description and mapping. Authority data, vocabularies and taxonomies therefore bridge dynamic linguistic and terminological variations. 4Memory distinguishes among universally accepted authority-data standards (e.g. GND, VIAF), subject-specific authority data

used by larger communities (e.g. Getty thesauri, KldB) and vocabularies and categories created by projects to answer their research questions.

Dual strategy for metadata harmonization and data contextualization

4Memory pursues a dual strategy to address the challenge of harmonising metadata in an international context: first, as explained in 2.2, we will use the existing memberships of our co-applicant and participating institutions in **national and international committees and standardisation initiatives** to develop existing standards of descriptive cataloguing and refine **international cross-disciplinary and cross-institutional data models** (e.g. EUROPEANA data model) regarding research data and the needs of historical scholarship and memory institutions.[12] Interoperable metadata models are crucial for the potential inclusion of data-holding institutions in data pools such as the 4Memory Data Space. These international activities are aligned with collaborations at the national level and in the NFDI's cross-cutting section "Metadata, Terminologies, Provenance", particularly with other humanities consortia.

Second, the diversity of historical data and its subjects means no one standard can describe all potentially relevant data without forfeiting information vital to its interpretation and reusability. The centralised approach of harmonising metadata models is thus complemented by advancing decentralised, research-driven practices. Integrating historical perspectives in authority data encourages a broader use of vocabularies and reference systems, also accounting for multilingualism, leading to interoperability with international standards (e.g. VIAF, HISCO, various Getty indexes). Standardised technical models (e.g. OWL and RDF) ensure **general interpretability without losing contextual information**. Aligned with the standards of memory institutions, these practices can contribute to a knowledge organisation system about data, building on minimum requirements for meta information while enabling automated exchange, international linking and modular extension. We pursue interoperability at a higher level of abstraction by identifying and mapping common concepts and entities in data and metadata.

Ontologies

In information science, an ontology is a shared, formalised, explicit and machine-interpretable representation of concepts (or terms) and their relationships to each other (Staab and Studer 2009). Ontologies model and represent knowledge, enabling the embedding of new insights into a knowledge system. Domain ontologies usually cover specific knowledge domains and evolve independently of each other: their representation is based on mathematical logic (description logics) and thus language independent. By mapping the applied concepts to controlled vocabularies (authority data), they can be addressed via authority data identifiers to help

[12] See Chapter 2 for international committees on metadata. Relevant data models include but are not limited to EAD for archives, LIDO / CIDOC-CRM for museums, MARC for libraries, and METS/MODS. See Appendix 1.2. for references of metadata and standards.

distinguish ambiguous concepts and allow concepts (and thus ontologies) to be maintained in several languages via labels. This is vital for historical disciplines and addresses the special needs of, e.g., area studies or ancient history. Identifiers also help identify common concepts in different ontologies, thereby interconnecting them. Linked ontologies form an important basis for knowledge graphs and Linked (Open) Data, allowing for a distributed development and maintenance of knowledge domains while ensuring interconnectedness and interoperability.

Knowledge Graphs

Knowledge graphs (KG) are based on ontologies (Singhal 2012). They supplement the concepts modelled in an ontology with references to concrete instances that represent them, e.g. object records in archives, museums and research-data repositories (Auer et al. 2018). Ontologies and KGs are thus suitable means for bringing together heterogeneous data sets and knowledge domains to make them interoperable, meeting the NFDI's goal of linking research data across disciplines. The four humanities consortia NFDI4Culture, NFDI4Objects, Text+ and NFDI4Memory will work closely together, with each covering the domain-specific requirements of their respective communities while sharing and reusing particular concepts when possible. 4Memory is also constructing an ontology and the associated knowledge graph in close contact with an international expert community, for example in dialogue with the Data4History consortium. In line with the requirements of its own community, we will focus on means of representing historical events; fuzzy/non-validated/contradictory facts; hypotheses; networks of people; and temporal, cultural and/or spatial variations in the meanings or labels of concepts. Key concepts will be developed by the consortium and extended through dialogue with our community, forming an approach with both centralised and decentralised aspects.

4.2.3 Approaches towards standardisation

Taking the points above into account, 4Memory addresses the following central challenges:

- Formalising community-specific, research-driven documentation practices into quality-assured metadata standards, leading to discipline-specific extensions of international established cross-disciplinary and cross-institutional data models (TA1);
- Enhancing the connectivity and discovery of research data through knowledge organisation systems with semantically clear and syntactically interoperable machine-readable metadata schemes and data models (TA2);
- Offering methods and tools for enrichment with and curation of authority data, vocabularies and taxonomies including automated entity recognition (TA2);
- Developing research data retrieval systems across information systems and metadata schemes with implementations of ontological concepts and/or indexing layers (TA3);
- Integrating the use of standards, quality assessment and data modelling into the basic skill sets of researchers and their methodology (TA4);

- Initiating a cultural change towards open data among researchers, their organisations and memory institutions through awareness raising and innovation (T5); and
- Fostering the reuse of research data across disciplinary boundaries and institutional structures through a moderated dialogue and common activities (TA6).

4.3 Implementation of the FAIR principles and data quality assurance

4.3.1 Obstacles, opportunities and concepts for data sharing

The FAIR principles (Wilkinson et al. 2016) aim to enable the provision and reuse of research data. These goals are met to varying degrees in different disciplines: in our community, many barriers to achieving them remain. Commonly noted obstacles to sharing research data are the additional effort and resources necessary to prepare them (Borgman 2012). In the historically oriented humanities, however, the time-consuming preparation of sources and research data is a given. In fact, the data types created in historical research and through the application of its methods – e.g. validation, provenance, transcriptions, editions, entity extraction, registries, contextualisation etc. – would directly improve the reusability of the sources in other research contexts. Researchers do not so much shy away from data *enrichment* but from its *publication* and final *revision*: they lack resources in terms of skills, support and funding and face a poor return on investment in terms of reputation and visibility in the scholarly discourse (Lemaire 2018, 238). 4Memory addresses the following obstacles relevant to our community and expressed in the problem stories collected during the application process (see 3.1.2):

A lack of recommendations and guidelines leaves researchers without orientation on how to find and adhere to appropriate research-data standards, even though – as many problem stories note – a general awareness of their necessity is already widespread. History remains one of the few humanities disciplines without domain-specific recommendations on research data management from the German Research Foundation (DFG). Providing guidelines and recommendations is a recurring element in 4Memory, as ***G**enerating standards for historical research data and sustainability* is a key objective. Their development through a participatory process is featured in several measures (TA1, TA2, TA3, TA4).

A lack of interoperability is directly related to the use or applicability of standards. The increased use of authority data and taxonomies (TA2), the development of a comprehensive ontology (TA1-3) and the creation of a knowledge graph (TA3) will address the challenge of (meta-)data heterogeneity and establish interoperability at a conceptual level. Moreover, the work programme implements collaboration among research, memory institutions and infrastructures, addressing the issue of interoperability at an organisational level (all TAs).

A lack of training in digital literacy leaves researchers without the basic IT and RDM skills they need to digitally enhance their workflows, gain skills in and perspectives on the analysis of digital data and extend their methodology through digital source criticism. The demand for postgraduate

training is currently unmet, and guiding concepts for skill profiles, curricula and learning material – for students, graduates, post-graduates or teachers – are still lacking, requiring discussion and efforts at development (RfII 2020b). Thus, support for existing international approaches is vital.[13] Anchoring a discipline-specific implementation of the FAIR principles in curricula and courses of study introducing digital methods is part of a wider cultural change process (TA4, TA5) enabling ***E**ducation and citizen participation* in humanities research.

A lack of a clear mandate for RDM allows researchers to shy away from data publications, which currently offer little professional reputation (e.g. in the context of granting tenure) and remain unintegrated in scholarly communication and quality management processes. 4Memory addresses these issues on multiple levels, creating incentives for researchers and providing support for research data publication. A reputation economy for the humanities must include publication processes for creation and review (RfII 2016, 46), a crucial point of intervention (Buddenbohm et al. 2016; Klump 2017; Cremer, Klaffki, and Steyer 2019) that TA1 and TA5 address; above all, we initiate, moderate and advance discussions on research data towards a new disciplinary data culture (TA1, TA4, TA5).

A lack of institutional support deprives researchers of needed RDM services. As temporary, third-party funded projects cannot compensate for the lack of institutionalised RDM support, we opt for different approaches to capacity building, such as developing knowledge exchange, consulting services, workflow recommendations and adaptable guidelines (TA1, TA2, TA4, TA5).

The many shortcomings in RDM mean that very little FAIR research data is currently available. We have set ourselves the goal of changing this situation and advancing data FAIRness as practice in the historically oriented humanities through the measures in all our TAs.

4.3.2 Mapping the FAIR principles onto the Task Areas

Findability

F1 : (meta)data are assigned a globally unique and persistent identifier
F2 : data are described with rich metadata
F3 : metadata clearly and explicitly include the identifier of the data they describe
F4 : (meta)data are registered or indexed in a searchable resource

Reliably identifying analogue artifacts is a longstanding practice in memory institutions. The use of archive, accession or inventory numbers as references is, similarly, a prevailing standard in historical research. However, many institutions and even established repositories still fail to provide persistent identifiers (PID) for digital resources. Working with the providers of relevant

[13] While generic RDM training materials are available more broadly and platforms for OER in the humanities have been launched (e.g. OER commons or DARIAH-CAMPUS), discipline-specific materials from the historically oriented humanities are rare, aside from pioneering approaches, such as the journal “programming historian” (digital history methods) and the platform “Ranke.2” (digital source criticism).

digital repositories and data collections, 4Memory will ensure that integrating PIDs – such as DOIs (Datacite) and URNs (DNB) – and providing services to assign and resolve identifiers are part of 4Memory's "Key Services" (F1 : TA3)[14]. Our concepts for retrieval, such as a knowledge graph and data bibliography, will always state the respective data identifiers (F3 : TA2, TA3). As rich description is essential to historical research, we seek to enhance its translation into public, machine-readable form. The selection of subject-specific topics describing data content – and enabling keyword-based queries – even without a common metadata schema is tackled through authority data identifiers, ontologies and a knowledge graph (F2 : TA1, TA2, TA3).

Datasets are registered in multiple indexes, such as bibliographic records of historical research data and its cataloguing as well as the aggregated search in the Data Space. In addition, other aggregation services such as OpenAIRE, BASE or Google Dataset Search are connected (F4 : TA2, TA3). While cataloguing builds upon established heuristic search practices, concepts like the knowledge graph introduce new retrieval methods and differ from traditional browsing and searching for sources in memory institutions' information systems. As a consequence, and to ensure findability, the heuristics for datasets (and their different logics) have to be integrated into historical methodology (TA4, TA5). The current lack of connectivity among the main scholarly infrastructures in our community (i.e. archives, museums, libraries, data repositories and data centres) has resulted in a fragmented scholarly record, leaving the digitised source at the archive, its transcription in a repository and a monograph at the library. Measures to increase connectivity and, thus, preserve the contextualisation of the data essential to the historical method also greatly increase the data's findability (TA2, TA3).

Accessibility

A1 : (meta)data are retrievable by their identifier using a standardised communications protocol
A1.1 : the protocol is open, free and universally implementable
A1.2 : the protocol allows for an authentication and authorisation procedure, where necessary
A2 : metadata are accessible, even when the data are no longer available

Accessibility is ensured by unifying technical interfaces for data access in the Data Space. The interfaces will build upon widely used and proven standards like REST and HTTP (A1, A1.1 : TA3, TA6). As not all resources are freely available (whether for data protection, third-party copyright or ethical reasons), the integration of authentication and authorisation mechanisms (A1.2 : TA3, TA6) will be addressed and discussed as a cross-cutting topic, such as in the NFDI section "Common Infrastructures". Together with the operators of the data collections integrated in the Data Space, we will ensure that the metadata for deleted or blocked objects remain available at least as a "tombstone record" (A2 : TA2, TA3). Apart from technical measures to ensure access

[14] Through this reference scheme, we explain *where* in our consortium these principles are addressed. In this example, principle F1 is particularly addressed by TA3. In many cases, multiple TAs will be named.

to data, enabling the publication of research data types (TA1, TA5) such as finding aids, registries and the description of previously unregistered sources – which are almost invariably produced by historical researchers as part of their standard research practices – will itself increase the accessibility of the analogue sources.

Interoperability

I1 : (meta)data use a formal, accessible, shared and broadly applicable language for knowledge representation
I2 : (meta)data use vocabularies that follow FAIR principles
I3 : (meta)data include qualified references to other (meta)data

Interoperability requires prerequisites at which memory institutions excel: precise descriptions, persistence and authority data. Interoperable metadata and data formats as well as the documentation of content and context are essential to data quality. Thus, we undertake a systematic survey, formalisation and representation of data-quality practices (I1 : TA1); of the use of vocabularies and authority files in our community (I2 : TA2); and of memory institutions' existing cataloguing and documentation rules (I3 : TA2). We aim not for a common (minimum) data-quality requirement for interoperability but instead develop means for data creators and users to assess data quality through evaluation criteria and defined quality levels (I1, I2 : TA1).

Controlled vocabularies, authority data and standards are vital prerequisites, but similarities and differences across time are key for historical disciplines. Temporal and cultural binding of terms is often related to specific entities, so historical and cultural perspectives must be integrated in authority files or added as their extension. We provide models for taxonomies and directories, and curation tools and automatable approaches allow greater semantic interoperability (TA2). Linking semantic information also reveals concordances among source representations (variants), authority data (persistent lemmas) and taxonomies (dynamic concepts) and interlinks the fields of museums/archives (variants/objects), libraries (authority data), research (dynamic concepts, taxonomies) and artificial intelligence (ontological concepts).

Reusability

R1 : meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1 : (meta)data are released with a clear and accessible data usage license
R1.2 : (meta)data are associated with detailed provenance
R1.3 : (meta)data meet domain-relevant community standards

The documentation of data is essential for other researchers to assure its quality and assess its value for further research. In the historically oriented humanities, rich descriptions enable the reuse of published research data: the data again become subject to scholarly analysis but within a different context, ultimately resulting in new perspectives and knowledge – hermeneutic

reusability means re-interpretability. Guidelines, assessment methods and tools for data quality selected or developed in 4Memory generally include rich metadata description as a quality feature (R1 : TA1, TA2). Requirements for data reuse are analysed according to domain-specific data types, the workflows and practices of the different communities and traditional referencing systems to provide formal quality criteria (R1.3 : TA1, TA2). Not all data necessarily require rich annotation to be reused, but their quality level must be transparent. Scalable quality criteria can be implemented in research settings from student exercises to large collaborative research projects (R1.3 : TA1). Provenance is a key concept of historical research, and digital source criticism must know the origin, genesis and modifications of data objects. We address the standardisation of gathering, compiling, storing and providing these metadata in several measures and the NFDI-section “Metadata, Terminologies, Provenance” (R1.2 : all TAs).

Memory institutions are increasingly opting to openly provide their metadata. In this respect, 4Memory builds on the preliminary work of the German Digital Library, which has already concluded data transfer agreements with almost 500 memory institutions regulating the subsequent use of metadata and, in most cases, data under open licences. As a result, the awareness of such licences and their impact has increased significantly. We aim to make the various data collections in the Data Space available under open and widely used licenses, preferably from the family of Creative Commons licenses (R1.1 : TA3). For reproductions, representations and sources, however, copyright issues hinder data publication significantly, both because of strict German and EU legislation but also a large amount of uncertainty among researchers and a lack of counselling (bwFDM 2015). We address the community’s need for guidance in legal issues and licensing directly with multiple tasks (TA5, TA6) and as a cross-cutting issue in the NFDI-section “Ethical & Legal Aspects”. In addition, political advocacy within (and through) the NFDI for improvements in copyright law is an important long-term effort.

4.3.3 Critical and domain-specific extension of the FAIR data principles

The FAIR data principles are necessary guidelines for advancing data sharing and data-driven research. They are, however, a starting point and mission statement rather than a complete solution: they do not address all issues of responsible data use arising in humanities research, especially with regard to physical research objects, research participants and people affected by the research or data publication. Thus, we aim to critically reflect on these gaps and extend the FAIR principles when implementing them into our domain in two crucial areas, explained below.

Reflecting on research ethics

Digital technologies and methods have increased the ethical challenges of historical research. The implications of narratives from the past for the future and an ethical responsibility to the dead are hardly new considerations, but researchers are now faced with access to a vast amount of human artifacts – through digitising the past and internet archives from the present – and the

software to investigate it in depth, including machine learning and artificial intelligence. Data providers are proliferating outside of memory institutions: traditional actors, “who have negotiated the ethics of historical data collection and use, such as librarians and archivists, are increasingly being sidelined” (Lin et al 2020). Memory institutions and historians are also reconsidering their practices from a postcolonial perspective, raising issues of restoration and the rewriting of history (Göttsche 2019). We will adopt the CARE principles into our global perspective on history, including non-European regions, particularly the Global South.

As outlined in Section 2, we face these ethical challenges through a multi-faceted strategy addressing different perspectives (juridical, cultural, ethical), various contexts (personal rights and privacy; the right to be remembered / forgotten; involvement and inclusion; the impact and social responsibility of research), tailored formats (discourse and publications, information material and training, services and recommendations, public statements and engagement). Fundamental to historical RDM, ethical questions are dealt with in all TAs and many measures, each with distinct foci: as part of quality assurance and assessment (TA1), integrated into documentation and authority data (TA2), operationalised in rights management in services (TA3), imparted through education (TA4), introduced into the scientific discourse (TA5) and built into consortial and disciplinary governance structures and society (TA6).

Curating diversity and complexity

Historical research thrives on dynamic rearrangements, differentiated knowledge systems and accounting for multiple perspectives. To unleash its potential and reach significant conclusions, humanities research addresses the variety of contrasting perspectives that critical reconsiderations of knowledge production have stressed in the wake of long-running historiographical debates and methodological disputes. Making information interoperable means interlinking systems that require stable persistence with the dynamic multiperspectivity of research data (and metadata). Scholars need to work with existing standards for describing and documenting particular kinds of research data but must also extend them to account for historical perspectives such as integrating models of uncertainty and change (Kuczera, Wübbena, and Kollatz 2019). Alongside technological concepts of harmonisation and interoperability, humanities concepts and priorities, such as diversity, multilinguality, historical context and complexity have to be preserved and introduced into computational and algorithm-based approaches to avoid the “risk of losing thick description” (Tóth-Czifra 2019) and a “future of record scarcity” (Rosenzweig 2003). Finally, the accessibility and reusability of datasets may impact the scope of research. We thus seek to advance the digital interface to sources that are protected or remain in analogue form, ensuring they are not overlooked or displaced by datasets that are simply easier to access and reuse.

4.4 Services provided by the consortium

4.4.1 Service Portfolio

The 4Memory consortium plans a comprehensive portfolio of services that build on each other – from data infrastructure to counselling and competence-building services – to achieve a cultural change in our community and to implement the LINKAGE objectives.

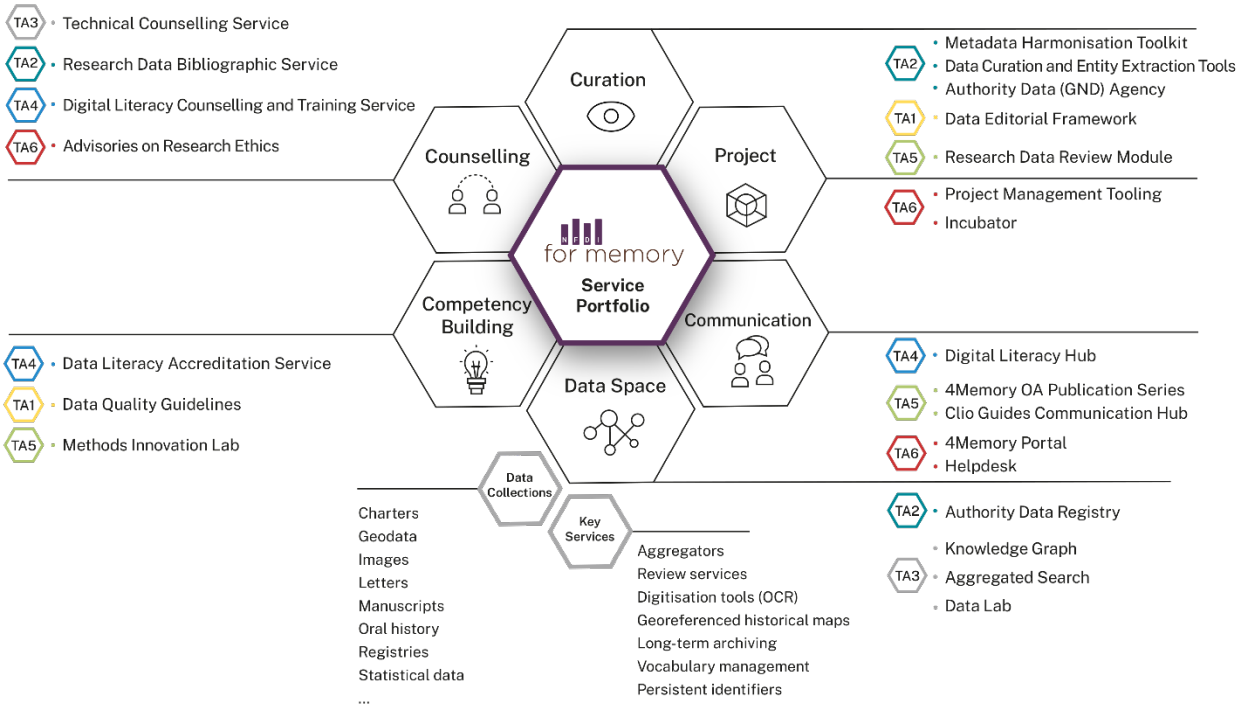

Figure 3: 4Memory service portfolio

Among its many benefits across our community, our portfolio provides many services for individual researchers, e.g.: a digital peer-review process that sets standards and ensures data quality (TA5-M4); the Data Space (TA3-M3) with its key services; the Labs for testing innovative DH tools and methods (TA3-M5 and TA5-M2); data quality guidelines (TA1-M4); tools for data curation, harmonisation and analysis (TA2-M3); data publication services (TA1, TA5) and various counselling services, including advisories on research ethics (all TAs).

The Data Space provides the core of the **4Memory Data Infrastructure**. It consists of key services covering important functional aspects, such as the persistent identification of resources (data objects), their secure and reliable storage (including long-term archiving) and management of authority data (vocabularies). These are supplemented by tools for digitising handwritten and non-Western sources (OCR) and handling georeferenced historical maps.[15] Finally, review and

[15] The resources of maps from the consortium highlight both the potentials and challenges of interoperable services on geohistorical data: Virtuelles Kartenforum (SLUB), Kartenspeicher (VZG), IEG Maps, Arcinsys (Archival geoinformation system for Hesse, HLA), Sammlung Perthes Gotha, Map Collection of the HI, Historischer Atlas von Bayern (BSB),

information services and aggregators allow a first level of linkage for heterogeneous offerings. The key services are accompanied by several important data collections from memory institutions, Specialised Information Services and research-data repositories covering the range of data types and topics of our community.[16] The knowledge graph interconnects key services and data collections on a conceptual level. The aggregated search based on the knowledge graph enables discovery across the diverse data collections. Using standardised technical interfaces, different tools and processes can access data in the Data Space and add data to it. Finally, an authority data registry complements the services of the 4Memory data infrastructure to support researchers in their search for suitable vocabularies and authority data. The selection of key services and data collections is neither closed nor exhaustive, focusing on essential and exemplary functions and on specific collection epochs and types. The aim is to create a nucleus for a functioning Data Space to demonstrate the advantages for historically oriented disciplines. Over the course of the project, we will, together with the respective providers, integrate further offerings into the Data Space, broadening and improving the offerings and addressing community requirements. Key services and collections are provided by consortium members as part of their institutional mission. The authority data registry, knowledge graph and aggregated search will be established in the NFDI framework.

The use of digital methods and tools is not standardised in our community. It will be a challenge for us to accompany the community's digital transformation and provide the necessary skills and competences. To this end, we plan several measures, particularly in TA4 and TA5, which will be reflected in **services for competence building**. The focus is on data literacy, data quality and the critical reflection of data methodologies. Over time, "train the trainer" concepts and the formation of knowledge networks will enable the community to disseminate methodological skills. Services for competency building will be established in the NFDI framework. The services are complemented by two closely interlinked labs: the Data Lab (TA3) and the Methods Innovation Lab (TA5). These two labs provide scientists with data resources and workflow environments to easily learn, develop and test new DH methods.

To support the adoption of new services, processes and methods, we also offer **counselling services** (all TAs). These contribute to digital transformation by helping researchers in the selection, application and sustainable provision of digital tools and processes, ensuring interoperability and fostering their considered and correct application. We help researchers evaluate and identify research data collections suitable for reuse. We educate researchers in

Military maps in the CAMO records (DHI Moskau), RM.Net (History of Rhine–Meuse area) (UT). See Appendix 1.2. Digital Resources for more details.

[16] Digital collections from the consortium about Jewish history may exemplify the diversity: Judaica of GU-UB (Manuscripts, incunabula, newspapers, prints, catalogues), Key Documents of German-Jewish History (edition of sources), Picture Database on Jewish life (images), The jewish Hamburg (encyclopedia) and Hamburg's Jewish cemeteries (geohistorical), Corpus of sources on the history of the Jews in the late medieval empire (spatial, epochal), Bibliography on the history of the Jews in the Bernkastel-Wittlich district (regional). See Appendix 1.2. Digital Resources for more details.

RDM according to the FAIR and CARE principles, and we develop guidelines for digital historical source criticism. We provide support with legal issues and address ethical questions in an advisory on research ethics (TA6). We will help historians create research data in machine-readable forms; label them with controlled vocabularies; and document, record or store them in accord with international standards. All counselling services will be established in the NFDI framework.

Services for the quality assurance and curation of research data pose a challenge to all consortia, but especially to historical disciplines, where positions such as documentalists or data scientists are rare. We will establish a subject-specific Data Curation Agency (as a GND-Agency, TA2) in cooperation with the DNB (Balzer et al. 2019). Data curators in institutions and projects can be supported in checking and improving data quality before publication, enriching information and contextualising the content. These steps are the basis for later merging and reusing data. Basic tasks and standards for the community must first be systematised and common goals defined. The focus is initially on the meaningful use of authority data identifiers, for whose implementation descriptions, scaling models and tools are provided. We also offer a data editorial framework focused on data publications, both at the stages of creation (TA1) and review (TA5). All curatorial services are aligned with the framework of the NFDI (TA6).

Cultural change in the historically oriented humanities can only take place through community involvement in the analysis of needs, the design and implementation of services and the dissemination of new and improved offerings, making **communication services** essential. The 4Memory portal will play a pivotal role in this process, bundling all services and presenting them clearly. It will be complemented by a helpdesk to which community requests can be directed and which will ensure their queries are forwarded to the responsible task area or areas. Further platforms (Clio Guides communication hub, 4Memory Open Access publication series, research data review module at H-Soz-Kult) also facilitate discussion with the community. All communication services will be established in the NFDI framework.

Finally, TA6 provides **project tooling** to steer the work of the consortium. It incorporates tools for project management and internal communication, also enabling the involvement of the participants and the wider community, e.g. via ticket and workflow systems, wikis and the like. Within 4Memory's Innovation framework the Incubator (TA6-M3-T1) enables project-based support and development of innovative concepts and services from the community to complement the work programme.

4.4.2 Service operation, security and sustainability

The permanent and sustainable operation of 4Memory's services is ensured on several levels. Key services are to a great extent operated by large infrastructure and memory institutions, whose mission includes providing such services. In most cases, therefore, the services are provided as

in-kind contributions and do not depend on 4Memory funds. For other services, TA3 and TA6 will work together with providers to ensure a stable technical and organisational operating environment. This can be achieved by the systematic acquisition of third-party funds, by leveraging synergies among providers and by increased usage, which can lead to a different positioning of a service in a host institution. Transferring services to infrastructure institutions (e.g. university computer centres) is being considered as a further option.

Data security plays a key role in the design of the 4Memory Data Space. Most historical sources are freely available and raise neither data protection nor ethical questions, but there are enough exceptions that we need to strive for general solutions. Large collections are mostly located in memory institutions such as archives, which operate according to specific legal requirements and have already implemented personal data protection in the analogue world. Solutions for the digital world are being developed, such as digital reading rooms, in which researchers can view access-protected materials if they have a proven scholarly interest and adhere to strict rules. However, this high level of security requires extensive infrastructure and know-how that research projects and small collections often lack. Through the organisational and technical optimisation of key services and the role and rights management in TA3, we will overcome this deficit. Cooperation with other NFDI consortia will be key in implementing an authentication and authorisation infrastructure. This is a complex task which cannot be solved by a single consortium, making it a vital cross-cutting topic. Finally, data security also touches on digital preservation, essential for security and long-term availability. We will thus include providers of long-term archiving solutions in the envisioned 4Memory data infrastructure.

We will also ensure the **sustainability** of our services through sustainable design features (RfII 2020a), which are a necessary part of every service concept within 4Memory and will be centrally coordinated (TA6). The service portfolio follows a best-of-breed approach: for specific functional requirements the most suitable solutions are selected and integrated into the infrastructure. Throughout the duration of 4Memory, we will continuously review together with the community whether the portfolio still meets researchers' requirements or if adjustments are required. Our governance (see 3.4) provides structures for community involvement in the review and decision-making process. Replacing individual services would have little or no impact on the overall portfolio. Still, adding new services to the infrastructure is associated with costs (e.g. for adapting technical interfaces, mapping data to the knowledge graph and taking authority data into account). For this purpose, either funds from the incubator (TA6-M1-T3) can be accessed or our participants can submit third-party funding applications together with service providers to accomplish such integration. The sustainability of the services is also guaranteed by the participating institutions, which already operate services in central areas such as long-term archiving (RADAR, DIMAG, Rosetta). Services integrated in 4Memory (e.g. H-Soz-Kult, historicum.net, DDB, Archivportal-D,

museum4punkt0) belong to the applicants' core mission and are thus also secured in their operation beyond the project funding.

Our service portfolio has been designed in response to problem stories from the historically oriented disciplines, addressing the needs they reveal. The **participative process** through which we are shaping our data infrastructure and services will ensure its development is research-driven. As described, the FAIR principles are followed and critically extended, as the design of our services also reflects. The services range from awareness raising, data literacy, training and consultancy to interoperable tools and data collections. The portfolio thus supports the full spectrum of our LINKAGE objectives. A significant proportion of the consortium's work does not primarily offer concrete services but rather initiates and promotes transformative community processes, which are of equal importance for 4Memory as its services.

5 Work Programme

→ See the also the illustration of the 4Memory project design (Figure 1) on the cover page.

Table 5.0: Overview of task areas

Task Area	Measures	Responsible Co-spokesperson(s)
TA1: Data Quality	M1: Assessment of data-quality practices M2: Analysis of data-quality requirements M3: Development of a data-quality measurement system M4: Development of recommendations and guidelines M5: Development of editorial frameworks for data creation M6: Competence network: implementation of data quality	Prof. Dr. Peter Haslinger, HI, Marburg Prof. Dr. Helmuth Trischler, DM, München
TA2: Data Connectivity	M1: Development framework for historical authority data M2: Working group on integrating historical data into authority data M3: Data curation and quality measurement interfaces and tools M4: Metadata cataloguing and documentation of research data M5: Interoperable data linkage for the Data Space	Dr. Klaus Ceynowa, BSB, München Dr. Katrin Moeller, MLU-HistData, Halle
TA3: Data Services	M1: Standardised and interoperable APIs M2: Technical implementation of a 4Memory Knowledge Graph M3: Creation of a 4Memory Data Space M4: Technical counselling service on interoperability and reusability of data M5: 4Memory Data Lab M6: Design and development of a virtual indexing layer	Prof. Dr. Gerald Maier, LABW, Stuttgart Matthias Razum, FIZ, Karlsruhe
TA4: Data Literacy	M1: Data Literacy Training Lab M2: Historical data criticism and RDM M3: Digital area histories M4: Digital research counselling service	Prof. Dr. Ursula Lehmkuhl, UT, Trier Prof. Dr. Stefan Schmunk, h_da, Darmstadt
TA5: Data Culture	M1: Epistemologies, knowledge orders, disciplinary contexts M2: Methods Innovation Lab M3: Ethical and legal implications M4: Publication and review culture with regard to research data and data-driven research M5: Creating impact, shaping change: enabling critical discourse in our community	Prof. Dr. Torsten Hiltmann, HU, Berlin Prof. Dr. Eva Schlotheuber, VHD, Frankfurt a. M.
TA6: Participation and Steering	M1: Participation and steering architecture M2: Academic coordination M3: Innovation framework and engagement M4: Communication and community management M5: Administration and procedures	Prof. Dr. Johannes Paulmann, IEG, Mainz

5.1 Task Area 1: Data Quality

5.1.1 Description of the Task Area

TA1 assesses and advances the definition, measurement and implementation of research data quality in historical research, enabling further academic study and computational research through FAIR data. It addresses several LINKAGE objectives. Working groups and events ***L****ink research, memory institutions and infrastructures*. Elaborating methods for data quality ***I****ntegrates historical source criticism into data services*. Assessing and implementing data quality practices strengthens the ***N****etwork of historically oriented research communities*. Quality-assured data shapes the ***K****nowledge order for the digital future of the past*, and quality criteria (e.g. provenance and processing information) ***A****dvance the analogue / digital interface of historical source material and data*. Data-quality guidance ***G****enerates standards for historical research data and advancing sustainability*. TA1 deals with content-related, administrative (e.g. legal documentation), technical and ethical aspects of data quality: standards for research data in historical research are (further) developed and implemented jointly with other TAs, the participating research and infrastructure communities as well as international bodies and networks. The understanding of “data quality” in TA1 is oriented on the definition of research data in the 4Memory RDM Strategy in regard to both metadata and data.

High-quality research data require that implicit data or research-related knowledge has been made explicit, i.e. through metadata. To enable FAIR reuse, it should be clear whether the data originates from a trustworthy or documented source; whether its content is accurate, contextualised and/or can be linked to other data; whether the data has been subject to automated or manual post-processing; and what ethical and other challenges such as ambiguity and multilinguality as well as multiple historical temporalities of data and their spatial connectivities have to be identified and met. Subsequent use rights as well as ethical considerations for further data usage (e.g. CARE principles) must be clarified and documented, and data must be available in technical formats suitable for archiving (including the necessary metadata). These general criteria must be adapted to historical research. The interoperability of data from (sub)disciplines that have thus far worked in non- or less-standardised ways with data from (sub)disciplines with standardised data methods – such as social and economic history or historical climatology and history of medicine – has to be considered. Practicable solutions need to be found for the integration of structured, semi-structured and unstructured data. Therefore, data quality will be assessed according to:

1) intrinsic data quality (accuracy, credibility, objectivity and reputation);
2) contextual data quality (added value, relevance, completeness and amount of data);
3) representative data quality (interpretability, intelligibility, consistency and scarcity);
4) technical and formal criteria (completeness of data release);

5) traceability (documentation and citation forms of sources used);
6) criteria for the form of publication (self-publishing / publishing);
7) findability / transparency (FAIR principles);
8) legal principles (licensing / data protection);
9) ethical considerations (e.g. the applicability of CARE principles for historical data).

The measures and tasks address all phases of the data life cycle: from planning a research project, through data collection, data preparation and analysis, data publication, data archiving and data reuse to data curation after the project's end.

- In M1, a working group will survey and assess existing practices (i.e. procedures, standards, requirements) for ensuring research data quality. Issues shared with other consortia (e.g. NFDI4Culture, NFDI4Objects, Text+) will be explored thoroughly.
- In M2, a working group will analyse and systematise requirements for data quality, applying and adapting FAIR and CARE principles to the needs of historical research.
- In M3, an expert group will develop a data quality measurement system, focusing on data types commonly used in the community, such as biographical interviews, time series, provenance and historical maps.
- In M4, guidelines and recommendations based on findings and discussions from M1-M3 are developed with the national and international community in settings within and beyond 4Memory.
- M5 develops data editorial processes to facilitate research data publication and initiate exchange between actors in research, publishing and data infrastructures.
- M6 establishes a community network to develop strategies, share experiences and support implementations for improving data quality on the institutional, national and international level and thus guarantees the transnational applicability of all measures.

Negotiation processes for developing quality levels for different research data types will be implemented within the scholarly community, allowing levels to be further developed in the future or for other content-related data types. The RfII has specifically noted the lack of subject-specific criteria for research data quality (RfII 2020c, 95). Together with TA4 and TA5, corresponding negotiation processes will be established in TA1. The community beyond the consortium will be invited to join the activities of the working groups of M1 and M2 and/or to review the proposals of the expert groups of M3, M4 and M6.

5.1.2 Measures

Measure 1: Assessment of data-quality practices

The first step towards ensuring research data quality in historical research is assessing the status quo, allowing general guidelines and procedures for improving sustainable quality management to be developed. TA1 leads a working group to survey the needs and practices of – and available

services for – data quality in our community (T1). This effort includes different data types and all phases of the research data life cycle, considering existing guidelines and data-collection principles in RDM plans, methods, software, plausibility checks and data curation, provenance and ethical questions and data protection needs, data documentation, indexing with metadata, principles of standardised documentation, the use of metadata standards and normative vocabularies, the licensing and citation of data, and standards of scholarly practice. Together with TA4-M2, M1 also addresses the challenge of incorporating historical source criticism including hermeneutical and epistemological dimensions into data-quality measures (T3). Apart from best-practices, obstacles to compliance with quality measures must be understood. Based on continuously collected problem stories, a systematisation of the surveyed requirements will further our understanding of what improvements are needed in data-quality practices (T2). A tangible outcome of M1 will be a directory of existing technical recommendations, guidelines and standards of data processing for historical research (T4). (Related problem stories: PS 2, 3, 15, 18, 29, 30, 49, 59, 67, 70, 73, 76, 81, 83.[17])

TA1-M1-T1: Assessment of needs and practices on data quality: Identification and recording of current forms of quality assurance and documentation of data collection and processing (including ontologies) and of community needs regarding data reuse in different RDM phases through 4Memory problem stories, surveys conducted by the participants, knowledge exchange with international partners and literature reviews.

TA1-M1-T2: Systematisation of community quality requirements: Descriptive information gathered in TA1-M1-T1 is turned into a dataset of structured information, allowing items to be clustered and quantified. This systematisation includes an inventory of practices and formal descriptions of workflows within them.

TA1-M1-T3: Assessment of standards for digital source criticism: Assessment of pre-digital methods and standards for source criticism and their implementation in digital data quality management measures, in close cooperation with TA4.

TA1-M1-T4: Establishing an overview of data processing standards: Setting up a structured list of existing technical recommendations, guidelines and standards for data processing in historical research.

Contributing members: *Co-applicants*: BSB, DM, h_da (HeFDI), HI, LABW, MLU-HistData, UT-SDCH, VHD; *Participants*: AK-PF, ArchivSCH, DDK, DHI Moskau, GEI, OIB, SUB

M1 cooperates with communities and institutions with relevant day-to-day experience (BSB, h_da, LABW, MLU-HistData, UT, ArchivSCH, DDK, GEI, SUB). VHD represents historians in many contexts, acting as a voice for the research community. OIB, DHI Moskau and partner institutions

[17] The 4Memory Problem Stories are available here: https://4memory.de/problem-stories-overview.

in the Eastern European Studies network coordinated by HI and BSB contribute expertise in multilinguality. AK-PF is a crucial participant in M1 regarding handling incomplete or incoherent contexual information, especially in ethically sensitive fields, such as objects collected within a colonial context or confiscated during Nazi persecution.

Measure 2: Analysis of data-quality requirements

In M2, the data-quality status quo revealed in M1 enables analysing data requirements and evaluating actions for improving and ensuring their quality (T2). Within this context, M2 also aims to make ethical issues an integral part of data-quality analysis via the FAIR and CARE-principles (T3, T4) This aspect will be complemented by NFDI4Objects in their Task Area 5 (Storage, Access and Dissemination) which will allow us to take into account both historical and archaeological research perspectives. A planned working group includes co-applicants, participants and other national and international partners (e.g. ICOM CIDOC) as intermediaries and experts for specific data-quality challenges. Data-quality requirements are analysed across the broad range of data types used in our community, based on problems named by our community and solutions developed in recent and ongoing projects within it (T1). Also, data use and quality management issues in various contexts (editions, archives, collections, libraries, museums, memorial sites, etc.) are addressed. A catalogue of FAIR-compliant practices is systematically developed, and recommendations regarding the optimisation of workflows throughout the data life cycles are made (T3). Particular attention is given to data transparency and traceability (e.g. data citation, licensing and persistence) (TA2). (Related problem stories: PS 1, 2, 3, 6, 15, 18, 29, 30, 49, 59, 67, 70, 73, 76, 81, 83.)

TA1-M2-T1: Assessment of tensions between quality requirements and existing conditions: Problem-oriented evaluation of tensions between quality requirements and the existing practical, technical, organisational and legal conditions in the community.

TA1-M2-T2: Comprehensive analysis of data-quality criteria: Analysis of data quality criteria for all phases in the research data life-cycle using different use cases and research data types, resulting in discussion papers, including proposals for graded quality levels.

TA1-M2-T3: Concretisation of the FAIR principles: Domain-specific concretisation of FAIR principles for historical research through a catalogue of practices and requirements concerning data types, formats and collections and FAIR-compliant clearing procedures.

TA1-M2-T4: Application and adjustment of CARE principles for historical research: Critical assessment of CARE principles for historical research regarding interpretation of personal and culturally sensitive data incorporating methods of conceptual history.

Contributing members: *Co-applicants*: BSB, DM, h_da (HeFDI), HI, MLU-HistData, UT-SDCH; *Participants*: AK-PF, ArchivSCH, DDK (KONDA and follow-up projects), DHI Moskau, DNB, ExClu AM, MWF Delhi, GDA, GEI, OIB, SUB, THK.

To analyse data collected in M1, its core team (i.a. AK-PF, ArchivSCH, DDK) will form a working group with institutions such as DNB and GDA (representing libraries and archives) and h_da with HeFDI and UT-SDCH (experienced in RDM / data curation). M2 works closely with co-applicants from TA2 (BSB, MLU-HistData). The working group will be supported by THK, profiting from its information science perspective and expertise in quality criteria analysis, and by SUB with its department “Metadata and Data Conversion”. DHI Moskau, ExClu AM, MWF Delhi, OIB give international views on FAIR and CARE principles.

Measure 3: Development of a data-quality measurement system

M3 develops a measurement system for historical research data quality. An expert group representing various subcommunities assesses existing generic data-quality certification systems (whether project-based or institutionalised). An overview of existing quality schemes is systematised and synchronised through dialogue with the community (T1). On this basis, a measurement system distinguishing various quality levels is developed, summarising the data-quality evaluation criteria and facilitating their application (T2, T3). The quality levels will enable the self-assessment of research data and facilitate a dialogue among data providers, data collecting institutions and data reusers. Existing peer-review standards and data-quality processes are evaluated in cooperation with TA5. Improving quality levels will enable interoperability (TA2 and TA3) and may qualify data for peer-review and scholarly publication (see TA5-M4). (Related problem stories: PS 2, 26, 29, 30, 37, 76.)

TA1-M3-T1: Determination of measurable data-quality criteria: Clustering of quality criteria and the creation of quality levels for these criteria.

TA1-M3-T2: Modelling quality criteria along the research data life-cycle: Quality criteria will be described and defined for all phases of the data life-cycle and selected data types.

TA1-M3-T3: Implementation into a modular measurement system: Creation of a modular (and expandable) building block principle for the identification of different levels of data quality, inspired by the Creative Commons licenses, in feedback with the community.

Contributing members: *Co-applicants*: BSB, DM, FIZ, HI, MLU-HistData, UT; *Participants*: ArchivSCH, DDK, DNB, MWF Delhi, THK

Experience in aggregating data is crucial to a successful data-quality measurement system. Thus, DDK and DNB are involved as participants. THK and MWF Delhi contribute use cases and, together with ArchivSCH and UT, contribute to a criteria system to evaluate and measure data

quality so it is suitable for reuse. TA1 also works with TA2 (BSB, MLU-HistData) and TA3 (FIZ) to guarantee data quality and standards for the Data Space and knowledge graph.

Measure 4: Development of recommendations and guidelines

In close collaboration with international boards and committees, an expert group translates the findings of M1-M3 into recommendations, catalogues and guidelines for the creation, assessment and future certification of research data (T2). The aim is to establish a permanent collaboration with existing international quality-assurance initiatives (e.g. GoFair, RDA) (T4). To that end, the expert group establishes a board for developing recommendations (T1) and identifies exemplary projects and work routines in which data quality is crucial for research and steps for developing suitable infrastructures for historical research. With regard to systematically implementing new project-management practices to improve workflows, this measure may serve as orientation for the DFG Review Boards and provide subject-specific information for the implementation of the DFG “Code” (Deutsche Forschungsgemeinschaft 2019), other third-party funding agencies and international (European) initiatives, committees and boards like CERI and RDA. Furthermore, the data-quality requirements for the 4Memory Data Space are defined (T3). Tasks in this measure are carried out in collaboration with TA2, TA3, TA4 and TA5. (Related problem stories: PS 1, 2, 3, 6, 29, 76, 83.)

TA1-M4-T1: Establishing a board for recommendations: Designing and establishing a board and the development processes for the definition of recommendations and guidelines for historical research and their continuous iteration within the community.

TA1-M4-T2: Development of recommendations and guidelines for the research data management process: Includes publishing relevant documents and preparing a manual for researchers addressing crucial questions above mentioned.

TA1-M4-T3: Defining quality requirements for the data space: Defining quality criteria, data types, metadata formats and general requirements for the selection of data from the common Data Space; developing criteria for transfer to long-term archiving (beyond 10 years) in cooperation with TA3.

TA1-M4-T4: Participation in quality assurance initiatives: Members of the consortium will continue and intensify existing engagement in interdisciplinary and international quality assurance initiatives (e.g. RDA) delivering input for and impact from 4Memory’s results.

Contributing members: *Co-applicants*: BSB, DM, FIZ, HI, MLU-HistData, UT; *Participants*: ArchivSCH, DDK, DNB, DSM, LMU-ITGW, SUB, THK

M4’s expert group includes UT, guaranteeing close coordination with TA4 for integrating the guidelines in teaching/training at universities and research institutions. THK and DNB act as intermediaries, connecting the group to other initiatives on regional (fdm.nrw), national (DINI,

nestor), European (DARIAH-EU) and international levels (RDA). In M4, it is crucial to cover the range of historically oriented institutions, such as universities (LMU-ITGW), libraries (BSB, SUB), archives (ArchivSCH), museums (DSM), and research institutions (DDK, FIZ, MLU-HistData) and to address the whole community.

Measure 5: Development of editorial frameworks for data creation

M5 aims to significantly lower the barrier of time-consuming data preparation through efficient workflows and a concept for dividing labour through editorial processes within an (organisational) framework applicable to existing institutional settings. Institutions in the consortium that engage in both RDM and traditional publication processes will jointly develop, implement and iterate editorial workflows for publishing research data. These data-editing processes – e.g. formal quality checks, format transformations, explicit documentation or metadata enrichment prior to publication – are developed with relevant actors in the key fields of research support, data management and publishing. Targeting formal data quality at an early production stage, the editorial procedures prime the peer-review process, focused on methodology and content, to be established by TA5. A curated network brings together actors from RDM and publishing departments (T1). Sharing existing workflows and jointly identifying needs create prerequisites for a data editorial framework designed in a working group (T2). The exploratory and preparatory work results in a service concept for data-editorial processes at the contributing institutions (T3). M5 will allow the outcomes and concepts of M1-M4 to be reviewed in practice and, if necessary, adapted or extended. The collaboration of established editorial offices and newly installed RDM positions initiates institutional change. (Related problem stories: PS 13, 37, 83.)

TA1-M5-T1: Curation of a network of actors from publishing and RDM: The curated network brings together actors from the fields of publishing and RDM to share experience and workflows, identify common goals and collaborations, and further development needs.

TA1-M5-T2: Modelling editorial processes in data-publication workflows: A designated working group from the network (T1) collects transferable best practices and evaluates examples to create formal process descriptions and structured sets of workflows. In addition, international players such as data journals and also commercial publishers will be consulted.

TA1-M5-T3: Concepts of data editorial processes as a service: Results of M5 serve as blueprints in the participating publishing and RDM departments to establish institutionalised services. The working group then seeks scalable recommendations for other settings identified in the network, e.g. data centres, research institutes, libraries, publishers.

Contributing members: *Co-applicants*: DM, HI, HU, IEG, MLU-HistData, UT-SeS; *Participants*: ExClu AM, HAB, UzK-CCeH / UzK-DCH

M5 is co-led by IEG. DM, HI, ExClu AM and IEG have (international) expertise in publishing, research scholarships and RDM; each addresses the entire historical research life-cycle in different settings. UT-SeS brings expertise from its DIAMANT model. MLU-HistData represents the disciplinary data centres in the working group. The HAB adds a research library perspective and the only digital humanities journal in Germany (ZfdG). UzK contributes expertise on workflows for data publications at DCH (data centre) and CCeH (research centre). HU provides long-term expertise in community-based editorial processes.

Measure 6: Competence network: implementation of data quality

Crucial to all TA1's activities is a constant dialogue with the community and all other TAs for the implementation and dissemination of data-quality measures. The guidelines will therefore be designed to meet different target groups (T1). Particular attention will be given to smaller institutions, academic committees and learned societies as actors that structurally lack the capacities and infrastructure needed to develop and implement digital data-quality standards and practices at their institutions. Therefore, this measure provides RDM guidance and activities for data-quality training in coordination with TA4 ("train the trainer") (T2, T3). (Related problem stories: PS 13, 29, 35, 37, 49, 67, 76, 83, 85.)

TA1-M6-T1: Preparation of guidelines for the first contact with researchers: Guidelines and guidance through the research data process.

TA1-M6-T2: Competence network for RDM capacity building: An open network for persons with RDM responsibilities is moderated by RDM experts from 4Memory to foster knowledge transfer and capacity building (transferred to practice in TA5-M5).

TA1-M6-T3: Promotion of data-quality implementations: Within the network, and in corresponding tasks in TA4 and TA5, the outcomes and deliverables of M1-M5 are presented and pilot implementations and practical experiences are discussed. Research institutions are offered an audit workshop based on the DIAMANT reference model.

Contributing members: *Co-applicants*: DM, HI, HU, IEG, MLU-HistData, UT, VHD; *Participants*: DDK, DSM, GDA, OIB

M6 brings together a group of intermediaries as well as MLU-HistData (TA2), UT (TA4) and HU, VHD (TA5). DDK contributes as expert in active dissemination of metadata standards in the field, GDA in coordination and communication with other data-providing institutes. OIB widens outreach on an international level. At the DSM a scientific assistant for digital documentation establishes digital quality management as well as a network of researchers working with data quality standards and documentation.

5.1.3 Milestones and Deliverables

Overview table on major milestones (MS) and deliverables (D) and the scheduled completion (project month).

Type	ID	Title	Month
MS	1a	Overview of data quality requirements and practices	10
MS	1b	Concept for enabling digital source criticism with regard to data quality	24
D	**I**	**Description of data quality standards**	**18**
MS	2a	Systematisation of requirement criteria	20
MS	2b	Assessment of legal and ethical challenges	22
MS	2c	Yellow paper on concretisation of the FAIR principles for historical research	28
MS	2d	Yellow paper on use, application and adjustment of CARE principles for historical research	32
D	**II**	**White Paper of graded levels of data quality criteria**	**36**
MS	3a	Concept of data quality measurement system along the research data life cycle	27
MS	3b	Concept of data quality levels and exchange with the community	38
D	**III**	**Community-approved data quality measurement system**	**42**
MS	4a	Achieve contributions to international data quality assurance initiatives	46
MS	4b	First draft of descriptions of quality requirements for the 4Memory data space	26
D	**IV**	**Recommendations and guidelines for the RDM process**	**54**
MS	5a	First draft of an editorial framework and exchange with the community	38
MS	5b	Pilot services for institutional RDM support of data publication	50
D	**V**	**White Paper: Data editorial framework for FAIR data creation and publication**	**56**
MS	6a	Guidelines for the first contact with researchers	54
MS	6b	Constitution of competence network for capacity building	38
D	**VI**	**White Paper on data quality implementations**	**60**

5.1.4 Cooperation with other task areas and cross-area dependencies

TA1 offers the key foundation for NFDI4Memory's work, and coordination with all task areas is essential. Research data can be made interoperable and transferred to the Data Space (TA3) only when data-quality criteria have been developed and met. To connect data and create the necessary infrastructural conditions, cooperation with TA2 and TA3 is vital. Key topics such as working with and developing the FAIR and CARE principles as well as the historical approach to data ethics will enrich the NFDI as a whole, e.g. via the NFDI e.V.'s cross-cutting section "Ethical & Legal Aspects". TA1 will work closely with TA4, since TA1's results and findings will flow into academia; in return, results emerging in TA4 via dialogue with the community, teachers and students will iteratively feed back into TA1's development. In M5, TA1 cooperates directly with TA5, aligning editorial workflows for data quality at the pre-publication stage with quality assurance during review processes. TA1 cooperates with TA6 on conducting the surveys, making use of consortial communication channels and community management (M4). Many aspects of data quality, its evaluation and compliance with relevant requirements pose shared challenges for other NFDI consortia. Thus, 4Memory will take into account the preliminary findings of the already approved NFDI4Culture and Text+ and complement them with historical methodologies (e.g. source criticism, provenance or data ethics). In accordance with the Memorandum of Understanding with NFDI4Culture, Text+ and NFDI4Objects, important aspects of data quality will be continuously discussed, developed and, where possible, collaboratively harmonised.

5.1.5 Quality Management and Risk Analysis

TA1 relies on the commitment and the expertise of the participants as well as a constant exchange with the community. Risks thus lie in the need to continuously maintain dialogue with the participants and the community at large. To manage this risk, three full-time staff positions spread over two institutions will be responsible for the extensive coordination and communication in TA1. Participants have been selected to ensure that memory institutions, research institutions and universities are represented in all measures, avoiding the risks of omitting parts of the community or inadequately accounting for international discussions. The working groups enable community participation and gathering broader perspectives during the collection and analysis processes. The expert groups ensure efficient production of the deliverables, particularly recommendations and guidelines, which are subject to public comment and review. The board for recommendations provides additional quality assurance.

5.2 Task Area 2: Data Connectivity

5.2.1 Description of the Task Area

TA2 advances the linking and contextualising of research data in the historical domain to enable a coherent historical record in the digital age and provide the foundation for the planned 4Memory Data Space. It lays the ground for historically sensitive, unambiguous data descriptions and the integration of research-led historical categorisation. Furthermore, TA2 focuses on cataloguing, automatic recognition and contextualisation within larger data structures. Developing and improving authority data, taxonomies (M1, M2), metadata (M4, M5) and ontologies (M5) will conceptualise the 4Memory Knowledge Graph in a domain-specific way. TA2 will increase the ***L****inking of research, memory institutions and infrastructures* via a knowledge graph, adding fundamentally to the representation of a ***K****nowledge order for the digital future of the past*. Connecting sources, publications and research data ***A****dvances the analogue / digital interface of historical source material and data*, as both sources and data will be accessible in one data space (M5). Open, jointly developed metadata and authority files ***G****enerate standards for historical research data and sustainability* (M2). To achieve full connectivity, TA2 will consistently act in accord with relevant working groups across the NFDI and internationally.

Many subject-specific authority data, taxonomies and controlled vocabularies are used in historical research (see 4.2; 4.3). Archives, museums and libraries also have a plethora of different metadata schemas and rules to describe their holdings. Against the background of this fragmented landscape, one of TA2's main goals is to lay the foundations for data connectivity and interoperability in the described domain. For this purpose, scholarly-driven, analytic classification systems must be connected with the cataloguing and description systems of archives, libraries and museums. Furthermore, the rules and standards applied by researchers and information infrastructures should be harmonised community-wide and in constant exchange with international bodies for (meta)data harmonisation (M1, M2, M4, M5). Via a 4Memory ontology, scattered research datasets can be made connectable. This is the basis for the meaningful

aggregation of data in the 4Memory Data Space via a knowledge graph (TA3). In this way, research data will not only become FAIR but will also be enriched with diverse semantic links and provide high analytic value.

In TA2, subject-specific authority data, taxonomies and controlled vocabularies are used to connect research data via their semantic content and allow historical phenomena to be described unambiguously as entities in human- and machine-readable form. Persistent and unique identifiers make heterogeneous data deriving from different research contexts and stored by various providers findable, comparable, analysable and reusable in a single knowledge space. By undertaking a research-led classification effort, M2 and M5 will offer terms to link denominations of entities with regionally different notations, whether within Germany, across Europe or worldwide. Current discovery tools rely on describing objects with structured metadata: information can be exchanged easily if providers use an interoperable metadata set, which requires harmonisation (M4, M5). Data with different metadata schemas can be brought together through integrating ontologies, necessitating that concepts, properties and relationships are defined in a data model (M5).

The contextualisation of knowledge within larger data structures relies on entity extraction and recognition. The use cases for authority data vocabularies described below (M1, M2: use cases 1-4) are the prerequisite for further development of innovative methods of automatic entity recognition and data curation services (M3). As writing history is not only about conducting research and publishing data but also finding and using research results (see 4.2.), another of TA2's key goals is integrating research data into the subject bibliographies widely used in our community (M4). In future, research data will be documented as independent publications and, if applicable, linked to research syntheses, increasing the data's visibility.

5.2.2 Measures

Measure 1: Development framework for historical authority data

M1 analyses existing authority data approaches for the long-term development of graduated concepts for search strategies, data curation, data exploitation and quality requirements. This applies both to the subject-specific context of 4Memory and to interdisciplinary and international initiatives. In doing so, 4Memory takes into account authority data that contextualise historical entities in a long-term perspective. TA2 begins by working with four use cases (UC) selected through community meetings and problem stories, bringing together diverse actors from 4Memory's international network. They cover concepts in different historical periods (UC1 Premodern, UC2 Modern/Contemporary History), historical geographical categories (UC3 Space) and personal relationships/activities and occupations (UC4 Activities). To this end, existing approaches are gathered and screened (T1), analysed and modelled (T3), as well as represented and described in a directory (T4) of subject-specific knowledge organisation systems, such as

authority data and controlled vocabularies. At the same time, existing activities are networked more closely in order to develop common goals and guidelines in the field of authority data production (T2, T5), this applies to specific concepts (UC 4), especially for a more in-depth cooperation with NFDI4Objects and Berd@NFDI.

TA2-M1-T1: Collection of overarching and suitable authority data and taxonomies: To support community-based authority-data management, we collect key international and national comprehensive and subject-specific vocabularies and authority-data standards and describe and evaluate them according to defined criteria (see PS 49).

TA2-M1-T2: Guidelines for the normalisation and indexing of research data: Based on use cases, guidelines for normalisation of data standardisation are discussed with the communities and best-practice examples for deriving rules are analysed. The focus lies on the contextualisation of information (see PS 40, 58, 72).

TA2-M1-T3: Graduated models of knowledge organisation systems: By evaluating existing models and experiences, we elaborate theoretical concepts for the rule-based differentiation of a scaled model of standards, authority data, taxonomies, controlled vocabularies and classifications so that they can be formalised and introduced into linked data structures or interfaces in the Data Space and to international initiatives.

TA2-M1-T4: Subject-specific directory for knowledge organisation systems: A directory with criteria and guideline recommendations is provided to orient consulting and project-planning regarding authority data and taxonomies.

TA2-M1-T5: Accompanying consulting on implementation: Develop consulting services on guidelines, conceptualisation and technical standards for authority data and taxonomies (see PS 56, 83).

Contributing members: *Co-applicants*: BSB, h_da, HI, HU, LABW, MLU-HistData, UT; *Participants*: **General**: ArchivSCH, BA, digiCULT, DHI Moskau, DNB, HU-UB, KBL, VdA; **UC1:** AdW-Göttingen, AdW-Mainz, BSB (Handschriftenportal), DHI Rom, FSU MEPHisto, FSU-II, MGH, SBB-PK; **UC2:** FU-CeDiS, DHI London, DIPF, FUH-Archiv, IFZ, IGDJ, DHI Washington, MLU-IZEA, UL, ZZF; **UC3:** CompGen, DHI Paris, ISGV, IStG, SLUB, UR; **UC4:** CompGen, DDK, DHI Paris, DIPF, GSWG, ISGV, SBB-PK, UBM, UR, WA-VfS, ZBW

In subgroups devoted to each use case, the participants take part in surveys and in-depth interviews and provide information on their requirements and applications of authority data.

Measure 2: Working Group on integrating historical data into authority data

The international working group aims to integrate and link authority data to the GND and other standards. The concepts derive from selected subject-specific authority data provided as use cases from the participants (T1), in order to meet the needs of both compact, largely standardised

keywording and diversified processes of scholarly analysis with basically the same networked approaches and concepts (T2). The aim is establishing a two-tiered, linked network of persistently addressable authority data (GND, metadata, search) in relation to dynamically changing data (taxonomies, ontologies, context data) geared towards scholarly analysis (T3). This approach addresses an essential goal for producing FAIR data by maintaining data transparency, traceability and contextualisation from the (analogue or digital) source to the national/international authority data set: it is essential to integrating historical source criticism (contextualisation), developing standards and networking different process levels of RDM (i.e. data production, analysis, provision and reuse; T4).

TA2-M2-T1: Authority data use case analysis: The working group explores the extraction, analysis and selection of vocabularies or authority data regarding scalable mapping possibilities, based on concrete use cases provided by members of the use-case subgroups.

TA2-M2-T2: Generalisable guidelines and principles for authority data: Generalisable guidelines and principles, which identify tasks and set priorities according to the use cases, are formulated, enabling the expansion of the GND as a subject-specific standard and a connection to international standard data.

TA2-M2-T3: Conception and foundation of Authority Data/GND agency: Priorities and tasks of a subject- oriented authority data and curation agency are defined, and common workflows and work packages determined (see PS 42, 66, 93, 95).

TA2-M2-T4: Community requirements for the development of standards: Development of workflows and offerings in interaction with controlled vocabularies, authority data/taxonomies and international standards based on the use cases (e.g. GND or KIdB2010) (see PS 6, 10, 18, 21, 29, 38, 42, 74, 75, 93).

Contributing members: *Co-applicants*: BSB, h_da, HI, HU, LABW, MLU-HistData, UT; *Participants*: **General**: BA, digiCULT, DNB, HU-UB, KBL, UE-FZG; **Use Cases:** Participants of the subject-specific use cases (see M1).

In use-case subgroups, participants discuss and formulate guidelines for a scaled model of data standards , providing subject-specific data for developing authority data models and taxonomies.

Measure 3: Data curation and quality measurement interfaces and tools

M3 addresses the conception, planning and preparation/development of tools for data curation (T1), content quality measurement (T2) and the automated indexing or enrichment of research data (T3; e.g. the definition of knowledge graphs, preparation for the Semantic Web or contextualisation of information) and providing the groundwork for their realisation in accompanying projects and/or the second NFDI funding phase. Together with TA1, criteria of a graded quality model for the authority data concepts of specific historical sub-areas are

developed. Basic concepts for the normalisation and contextualisation of data as well as for its taxonomic preparation are subjected to quality control and prepared for transfer into tools and interfaces (TA1, TA3, TA4, TA5).

TA2-M3-T1: Tasks of data curation: Aiming for generalisable requirements for tools of data curation and the contextualisation of entities or for creating taxonomies, use-case participants will formulate the challenges they face and the procedures they require on the basis of precise problem analyses (see PS 1, 10, 23, 29, 54, 58, 89).

TA2-M3-T2: Concepts for the comprehensive use of authority data and tools of data curation: TA1's results will be generalised to develop comprehensive principles for subject-specific forms of quality control, plausibility tests and data curation. Data-specific features, such as subject-specific and technical data formats or entity types, are considered.

TA2-M3-T3: Automatable approaches for authority data enrichment: To enable the use of normative data in historical and original-language research and the quality measurement of data structures according to selected criteria, existing data-extraction procedures, data-enrichment methods and quality-improvement approaches will be tested and evaluated.

TA2-M3-T4: Derivation of recommendations for the quality assessment and certification of data, methods and tools of data curation: Together with TAs 1, 4 and 5, recommendations for quality measurement and the identification of modular quality principles will be derived from the findings of M3. In addition, data-specific, content-related principles will be established for a scaled model of quality principles (see PS 3, 38, 53, 84).

Contributing members: *Co-applicants*: BSB, h_da, HI, HU, LABW, MLU-HistData, UT; *Participants*: General: BA, digiCULT, DNB, HU-UB, KBL, UHH-FDM; Use Cases: Participants of the subject-specific use cases (see M1).

Participants provide data to create tools for data curation and make them available under open licenses. Expert workshops and discussions assess resulting curation tools, develop guidelines and define further requirements.

Measure 4: Metadata cataloguing and documentation of research data

After evaluating the status quo and its desiderata (T1), expert groups develop common standards and rules to describe and document research data (T2), easing data linkage and aggregation in the 4Memory community and its Data Space, as well as with international services. Scholars will have a shared set of descriptors, authority data and taxonomies to search for sources and research data. Service providers such as archives, libraries and museums will be given clear guidelines on cataloguing and documenting research data based on common standards and rules adapted to them (T3). The inclusion of a defined set of research data – related to publications and as publications in their own right – into the German Historical Bibliography (GHB) and possibly

other bibliographies[18] provides a proof of concept for recommended formal descriptors, authority data and taxonomies (T4). As described in the RDM-Strategy, 4Memory's research community will benefit from the extended GHB by being able to use this new resource in an established environment with direct-access options.

TA2-M4-T1: Evaluate existing rulesets for the description of research data: Gaps in existing national and international cataloguing and documentation rules in memory institutions and their use of authority data to describe research data are identified, discussed with experts and addressed with relevant governance bodies in this sector (see PS 53).

TA2-M4-T2: Toolkit for cataloguing and documentation of research data: A toolkit will be created and distributed amongst archives, libraries, museums and disciplinary repositories to foster the adoption of a common set of rules in order to achieve the harmonised description and documentation of research data in the 4Memory community (see PS 3, 59, 69, 81, 89).

TA2-M4-T3: Workflows for RDM in disciplinary repositories and GLAM collections: Workflows for descriptive cataloguing, documentation and subject indexing of research data will be conceptualised and provided to infrastructure and memory institutions.

TA2-M4-T4: Integrate research data into bibliographic information systems: Research data, defined by specific criteria and selected from different sources, will be brought into the GHB and other established information services like FID-Portals, subject bibliographies or union and international catalogues.

Contributing members: *Co-applicants*: BSB, DM, MLU-HistData, HI, UT; *Participants*: ArchivSCH, BA, digiCULT, DHI Rom, FB Gth, FSU-MEPHisto, DHI Washington, GDA, GNM, HAB (Handschriftenportal), IFZ, KBL, LMU-ITGW, OFU, SBB-PK (Handschriftenportal, Kalliope, Orient Digital), UE FZG, UK-UB, UL, UzK-CCeH, UzK-DCH

BSB coordinates all tasks, supported by co-applicants MLU-HistData and UT, with experience from a research perspective. All participants contribute their respective expertise on metadata on a national and international level; those from memory institutions, data repositories and research institutions especially on cataloguing, documenting and describing research data; those representing research (e.g. DHI Washington) provide sample projects as test cases for the work package outcomes. DHI Rom, IFZ and KBL also cooperate on bibliographies.

Measure 5: Interoperable data linkage for the Data Space

For the exchange, integration, delivery and reuse of research data, a 4Memory ontology and a metadata schema must be developed (T1, T2) by a metadata competence unit. This will allow linking the research data of the 4Memory community within our Data Space and connecting it to

[18] e.g. Gnomon Bibliographische Datenbank, ARTOS, regional, epochal or subject bibliographies bibliographies like Regesta Imperii, Bibliografischer Informationsdienst of IFZ etc.

other aggregation services such as BASE and Google Dataset Search or future services in NFDI and EOSC. The developed data model must also integrate non-German languages as well as knowledge and organisation systems in order to represent research data from fields such as area studies, to introduce postcolonial perspectives and to secure the international connectivity and visibility of data curated and listed by 4Memory, thus enabling a truly transnational historiography (T3). Additionally, persistent identifiers will allow addressing granular data, improving the citation and reuse of research data through, e.g. linking to micro-publications such as annotations or referencing concepts within taxonomies (T4).

TA2-M5-T1: Conceptualisation of the 4Memory ontology to aggregate research data: With TA1, an Ontology Expert Group conceptualises subject-specific aspects of the 4Memory ontology built by TA3 to enable data linkage and aggregation in a knowledge graph and in the 4Memory Data Space. The ontology will be compatible with other NFDI ontologies under development and international ontology initiatives in this field. The ontology will combine the representation of basic relations like creator or title with historically relevant concepts and relations referring to events and periods, persons, names and boundaries of specific places (which may alter over time), changing meanings of terms and notions, and uncertainty about factual data. As other consortia need the same concepts, e.g. provenance in the case of NFDI4Objects, cooperation in developing the definition is planned. A plan for the ontology's organisational structure will support subject-specific modular extension and adjustments as well as connectivity to other NFDI-consortia based on their ontologies (see PS 54).

TA2-M5-T2: Development of a 4Memory metadata schema for the Data Space: A 4Memory metadata schema will guide interested institutions, helping them to describe research data in line with the community's needs and to integrate them into the Data Space.

TA2-M5-T3: Handle multilingualism in research data description: Multilingual description of research data will be made possible by taking into account the special features of non-German languages (writing systems, coding etc.), the use of multilingual authority data and controlled vocabularies from across the world and options for automatic (text) analysis for in-depth subject indexing (see PS 40, 41, 42, 43, 47, 50).

TA2-M5-T4: Recommendations for the usage of persistent identifiers: Established persistent identifier systems will be assessed regarding their suitability for ensuring the long-term referenceability of research data. The assessment will be coordinated with respective initiatives of other NFDI consortia and result in recommendations for the application of persistent identifiers in a broad range of usage scenarios.

Contributing members: *Co-applicants*: BSB, DM, FIZ, HI, HU, MLU-HistData, UT; *Participants*: ArchivSCH, BAdW, DHI Moskau, DHI Rom, FSU-MEPHisto, FU-UB, DHI Washington, GNM, GU

UB, HAB (Handschriftenportal), KBL, LMU-ITGW, MGH, MPIWG, OFU, SBB-PK (FID CrossAsia, Handschriftenportal, Orient Digital, Kalliope), UK-UB, UL

BSB coordinates all tasks. Co-applicants DM, FIZ, HI, HU and MLU-Histdata are involved in developing the 4Memory ontology, UT contributes to multilingualism. Participants BAdW, GNM, HAB, FSU-MEPHisto, MPIWG, KBL, LMU-ITGW and OFU bring in their expertise in ontology development. Participants HAB, KBL, MPIWG SBB-PK (Handschriftenportal, Kalliope, Orient Digital) and UK-UB assist work on metadata schemas and harmonisation. Multilingualism is covered together with ArchivSCH, BAdW, DHI Rom, DHI Moskau, FU-UB, DHI Washington, GU-UB, MPIWG, OFU, SBB-PK (FID CrossAsia), UK-UB and UL. Sample data from research projects of FSU-MEPHisto and DHI Washington is used for feedback and as proof of concepts. KBL and MGH contribute expertise on PID and research data.

5.2.3 Milestones and Deliverables

Overview table on major milestones (MS) and deliverables (D) and the scheduled completion (project month).

Type	No	Title	Month
MS	1a	Formalised models of knowledge organisation systems	15
MS	1b	Guidelines for the standardisation of normalised data and vocabularies	26
D	**I**	**4Memory directory of knowledge organisation systems**	**26**
MS	2a	Concept for a GND agency	32
MS	2b	White Papers on Authority Data Use Cases	40
D	**II**	**Foundation of a GND-Agency**	**40**
MS	3a	White Papers on Data Curation (Use Cases)	53
MS	3b	Quality testing of tools for automatic data enrichment	60
D	**III**	**Recommendations for data quality assessment and certification**	**60**
MS	4a	First version of research data description toolkit finalised	36
MS	4b	Set of workflows for research data in repositories and reference systems	48
D	**IV**	**Cataloguing of selected research data in the GHB and other bibliographies**	**55**
MS	5a	Set of use cases and approaches for multilingual description of research data	48
MS	5b	Recommendations for the application of persistent identifiers	60
D	**V**	**Implementation concepts for 4Memory Ontology**	**48**
MS	6a	Concept for a 4Memory Metadata scheme	36

5.2.4 Cooperation with other task areas and cross-area dependencies

TA2 aims to link data and systems of knowledge organisation from scholarly research, memory institutions and information infrastructures, making inner-consortional cooperation central to its work programme. With TA4 and TA5, the use of common standards, rules and interfaces will be promoted by engaging with the discipline-specific research and publication culture. With TA4, competences and tools for using and generating metadata and authority files will be conveyed that are relevant to a range of users, from citizen scientists to international researchers. The data populating the Data Space will adhere to discipline-specific quality criteria to be drafted with TA1. Communication efforts are supported by the community management of TA6, while TA2 itself acts as a multiplier both for the memory institutions and its particular group of participants on a national level and within international bodies of metadata harmonisation and data standardisation.

TA2, with TA4, will ensure that multilingualism issues are respected consortium wide. One of 4Memory's main outcomes will be a subject-specific 4Memory Data Space, based on a knowledge graph, with the help of semantic web technologies. For this, TA2, with TA1, conceptualises the requirements to be technically implemented by TA3. The size and diversity of the challenges in the fields of authority data, metadata, ontology development and interoperability require that TA2 cooperates closely with other humanities consortia and with the sections "Common Infrastructure" and "Metadata, Terminologies and Provenance" in NFDI e.V.

5.2.5 Quality Management and Risk Analysis

TA2 pursues the ambitious goal of bringing together a large group of diverse partners from research institutions and the GLAM sector to collaboratively develop common standards, rules and authority data. The networking and feedback mechanisms and communication structures that TA2 will establish are essential to ensure the continuous involvement of the respective partners and guarantee community-wide implementation of the developed quality standards, ontologies and metadata schemes. Bridging existing differences in scholarly and institutional practices is often a time-consuming and communication-intensive effort, as is implementing new standards and procedures in memory institutions. All changes made with regard to established standards require the consent of the respective governance bodies. Hence TA2 has to address two major risks: the developed standards must fit the requirements of both research and memory institutions and need to be coordinated with similar efforts in other consortia. Additionally, given restricted budgets, implementing the developed quality standards and recommendations must be cost effective. TA2 will meet the latter challenge by always considering the cost-benefit balance of our recommendations.

5.3 Task Area 3: Data Services

5.3.1 Description of the Task Area

TA3 joins key services and data collections in a virtual 4Memory Data Space that allows historians to work with data in both traditional and innovative ways, enriching them and making them available again to the historically oriented humanities and beyond. It also offers a technical counselling service to empower community members to expand the Data Space with their own services. Over decades, researchers and memory institutions have applied numerous individual RDM solutions, leading to heterogeneous, single-use “data silos” that make comprehensive searches and discovery, at best, labour-intensive and often impossible, impeding the reuse of data and services. The elaborate indexing within projects has yet to flow back to the memory organisations providing the sources, a bitter loss of research output that remains unavailable to third parties. Researchers often even lack solutions for securely storing research data during and after their scholarly work. Addressing these challenges, TA3 pursues five goals:

(1) Make heterogeneous and distributed data collections and services from research and memory organisations FAIR and easily usable for researchers and algorithms;

(2) Improve interoperability and sustainability of project results by elaborating and publishing guidelines, best practices and standards;
(3) Enable innovative methods and processes in the field of digital humanities;
(4) Increase the amount of digital and machine-readable data sources as a prerequisite for these methods and processes;
(5) Conceptualise mechanisms to reuse project-based indexing efforts in memory institutions.

TA3 has six measures: M1 defines programming interfaces (APIs) for data collections and key services, while M2 lays technical foundations for the community-based construction of a knowledge graph. They are the basis of the 4Memory Data Space, in which services and data collections remain distributed but become comprehensively discoverable, searchable and accessible (M3). M4 gives guidance on bringing data and services into the Data Space. The Data Lab (M5) facilitates a technical platform for innovative new methods (working with TA5's Methods Innovation Lab) to easily assemble, analyse, annotate and enrich the Data Space's data with enhanced indexing information. M6 develops a concept to deliver such data back to the collections in the Data Space, thus increasing the depth of indexing and data quality.

The core of TA3 is the Data Space, consisting of both data sources and key services. From the problem stories (see 3.1.2), we extracted requirements from our community and identified appropriate key services that address these essential needs:

- the provision of reliable storage space and long-term archiving of research data (e.g. DIMAG of LABW and HLA, RADAR of FIZ);
- authority data servers with corresponding vocabularies and authority files (e.g. GND of the DNB, DANTE of the VZG or xTree of digiCULT);
- services for persistent identification (e.g. URN service of DNB or DataCite);
- tools and technologies for full-text recognition (e.g. OCR-D, Transkribus) and georeferencing (Virtuelles Kartenforum of SLUB, Geodaten-Repositorium of IStG);
- aggregators (e.g. German Digital Library or German Archives Portal) that provide instant access to various data collections and repositories provided by memory organisations, research institutes, universities, and projects; and
- information services for historical research (H-Soz-Kult, Clio-online, Historicum.net).

Integrating key services into the Data Space requires their FAIRification. Interoperable technical interfaces (M1), the consistent use of systems for persistent identification of data, and ensuring sustainability and permanence (M3), create a technical basis that must be complemented by interoperable data and metadata. Historical scholarship involves heterogeneous research questions, methods, tools and repositories, leading to numerous metadata schemata. This diversity cannot be replaced by a centralised approach but must be tackled conceptually. M2 implements a knowledge graph (KG) allowing aggregated search and discovery systems. We will

provide specifications to make tools and data collections compatible with the Data Space, enabling researchers to bring their collections and tools into it, making their findings more visible and reusable. M4 supports this process via a targeted technical counselling service to ease integration, enabling community-relevant data and tools to enter the Data Space. Many research projects in historical disciplines index sources from memory institutions in greater depth, according to their research questions. The 4Memory Data Lab leverages unified access to data and services to give researchers a platform to explore and apply new methods in machine learning, semantic annotation and advanced OCR techniques (M5). The in-depth indexing information created there (up to machine-readable full texts) is also generated in conventionally designed projects. In M6, TA3 designs a virtual indexing layer that combines, on a virtual level, indexing information from research projects with that of memory institutions, making it searchable and reusable. The virtual indexing layer will be based on the KG.

TA3 contributes to the LINKAGE objectives ***L****inking research, memory institutions and infrastructures*, ***I****ntegrating historical source criticism into data services*, ***K****nowledge order for the digital future of the past*, ***A****dvancing the analogue / digital interface of historical source material and data* and ***G****enerating standards for historical research data and sustainability*.

4Memory's service portfolio will enable historians to easily publish, archive, discover and retrieve FAIR data via the Data Space, opening up data silos, linking collections in research and memory institutions and providing interfaces between the digital and analogue worlds.

5.3.2 Measures

Measure 1: Standardised and interoperable APIs

Historians use heterogeneous data sources. The compilation of project-specific resource collections spanning multiple data sources is therefore a laborious intellectual process. Unified programming interfaces allow standardised, interoperable access to these heterogeneous data. M1 will therefore define such an interoperable API together with the providers of key data sources and foster their implementation. This is important preparatory work for the Data Space planned in M3. We will closely monitor international developments such as IIIF, SPARQL or GraphQL and coordinate with the other consortia of the MoU group as well as the "Common Infrastructure" section of the NFDI. M1 is organised in three tasks (related problem story PS 11):

TA3-M1-T1: Access API for heterogeneous data sources: This API provides unified, machine-readable access to research data and metadata from heterogeneous data sources, covering search and read functionality.

TA3-M1-T2: Data exchange API for technical interoperability: This API extends the access API by write (i.e. create and update) functionality. It enables data exchange between data sources as well as between research projects and memory institutions.

TA3-M1-T3: Viewer API for different research data types: Existing viewers for formats like visual or textual content, tabular data, geographic coordinates or 3D models will be evaluated and a selection implemented in the 4Memory Data Space. Recommendations will be made to DH-tool developers to include data types not yet covered by their projects (see PS 38, 63).

Contributing members: *Co-applicants*: BSB, FIZ, MLU-HistData, UT; *Participants*: BAdW, DHI Rom, DHI Washington, HAB, KBL, MPIWG, OFU, UHH-FDM, UL-UB

BSB coordinates all tasks. All co-applicants and participants provide experience with different APIs and viewer technologies to the 4Memory consortium and/or offer data and APIs for testing.

Measure 2: Technical implementation of a 4Memory Knowledge Graph

M2 builds the KG's technical foundation alongside a distributed editorial system enabling federated maintenance: various projects will be able to participate and contribute. Among other benefits, distributed editing gives non-European communities decisive influence over the representation of their cultures in knowledge systems, supporting the observance of CARE principles. The federated approach will raise complex questions of data quality that are addressed in TA1. The KG's core component is an ontology conceptualised in TA2 in line with international standards. It will cover specific concepts and knowledge domains relevant to historians, such as temporal fuzziness, the changing semantics of terms over time, hypotheses, and temporal and spatial networks. The KG enables the semantic linking of the heterogeneous data sources of the Data Space (M3), allowing for interlinking with external resources, e.g. from other NFDI consortia or internationally. All the KG's data will be queryable via a SPARQL endpoint. M2 thus undertakes three tasks (related problem stories: PS 6, 10, 20, 29, 54):

TA3-M2-T1: Establish technical foundation: Together with an expert group and in cooperation with the MoU consortia, FIZ will lay the technical foundation for a KG and a collaborative editorial environment for its distributed population and ongoing maintenance. FIZ and LABW will implement data curation guidelines developed in TA1 and TA2 (e.g., via SHEX), enabling quality-assured KG population by non-experts in accordance with the guidelines.

TA3-M2-T2: Populate the 4Memory Knowledge Graph: With broad involvement of the community and the consortium as a whole, LABW coordinates, with TA2, the population of the KG. The distributed editing environment enables the contribution of concepts from diverse research projects. This ongoing effort will be overseen by the established expert group.

TA3-M2-T3: Provide 4Memory KG resources as Linked (Open) Data: To allow third parties to perform complex analyses or to integrate the KG into their own systems, FIZ will provide access to the triples via SPARQL endpoint(s).

Contributing members: *Co-applicants*: FIZ, HI, HU, LABW, MLU-HistData; *Participants*: CompGen, digiCULT, DNB, ExClu AM, IFZ, MPIWG, UHH, VZG

All contributing members will engage in the expert group, contributing services (CompGen (GOV), digiCULT (xTree), VZG (DANTE)), technical expertise (FIZ, DNB, HI, MPIWG, UHH) or broad community knowledge which is necessary to populate the KG (HU, IFZ, LABW, MLU-HistData). ExClu AM ensures the connection to non-European/African perspectives.

Measure 3: Creation of a 4Memory Data Space

M3 designs and implements the 4Memory Data Space, a federation of heterogeneous data sources from research and memory organisations, which we connect via common interfaces (M1) and a Knowledge Graph (M2). This allows researchers to perform cross-searches and compile corpora from different sources. M3 accomplishes five tasks (related problem stories: PS 3, 11, 23, 29, 30, 31, 44, 70, 76, 81, 83):

TA3-M3-T1: Edit a technical specification ("blue paper"): A working group (FIZ, LABW, TA2 and participating service providers) develops and publishes a technical specification ("blue paper") for integrating data sources into the Data Space. Standards include APIs (M1), support for the knowledge graph (M2), data formats (e.g. METS/MODS, LIDO, EAD) and the adoption of indexing practices, e.g. by using recommended authority data. These steps contribute to the data FAIRness, building upon e.g. the German Archives Portal or German Digital Library.

TA3-M3-T2: Adapt selected key services to the 4Memory standards: FIZ supports the working group's members to adapt their services to 4Memory standards (as laid out in the blue paper) and integrate them into the Data Space. In this work, functional optimisations of key services are also carried out with regard to the annotation, transcription, semantic enrichment and georeferencing of research data. Other service providers can later integrate their offerings into the Data Space by implementing the blue paper's standards and interfaces.

TA3-M3-T3: Develop a discovery search service: FIZ and LABW will develop an aggregated search service for the Data Space based on the KG created in M2, cooperating with Measure 5.4 from NFDI4Objects. The KG will allow for navigating access to the research data and a semantic search across all resources.

TA3-M3-T4: Design a policy framework for rights management: Apart from legal restrictions (e.g. data protection and copyright), ethical aspects may prohibit access to data (see TA5-M3 and TA6-M2-T4). Requirements for text and data mining, such as machine-readable licenses, must also be considered. LABW and FIZ, with the community and coordinating with TA5 and TA6, will design a rights management policy framework in the Data Space based on international standards. We will cooperate with other NFDI consortia in the relevant NFDI section.

TA3-M3-T5: Provide services for long-term archiving of research data: TA3 will analyse the offerings of the participating infrastructure facilities for reliable storage and long-term archiving. Based on the results, TA3 will prepare a comprehensive offer for federated services and provide

standardised organisational and technical interfaces (e.g. to established systems like DIMAG or RADAR) to ensure professional long-term archiving of research data. This task will be accomplished in cooperation with Measure 5.1 of NFDI4Objects.

Contributing members: *Co-applicants*: BSB, DM, FIZ, HI, HU, LABW, MLU-HistData, UT; *Participants*: Clio-online, CompGen, digiCULT, DNB, ExClu AM, FU-CeDiS, HLA, HU-CMS, IFZ, IStG, LRZ, OFU, SLUB, SBB-PK, UHH, UE-FZG, UT-SEAL, VZG, ZDV

Contributing members form the working group. With TA2, digiCULT, FIZ, HI, HU, LABW, MLU-HistData, UT-SEAL, UE-FZG, UHH and ZDV agree on the Data Space's specifications and technical interfaces. Members integrate data resources: FIZ/LABW (Archivportal-D, DDB), IFZ (EHRI Portal), SBB-PK (FID CrossAsia, Handschriftenportal, Kalliope), BSB/DM (historicum.net), Clio-online (H-Soz-u-Kult/Clio-online), HU-CMS (Laudatio), HLA (Repositorium für Digitalisate), FU-CeDiS (Oral-History.Digital), IStG (Geodaten-Repositorium), SLUB (Virtuelles Kartenforum), ExClu AM (non-European/African repositories), UE-FZG (FactGrid), UT (eAQUA-Portal), VZG (Kartenspeicher, K10plus, DANTE). LRZ provides software development expertise. DNB and UHH will ensure interoperability with other NFDI consortia and networks. UT-SES/ZDV (iRODS/ViDa), FIZ (RADAR), LRZ (FDM services) and HLA/LABW (DIMAG) provide services for long-term archiving.

Measure 4: Technical counselling service on interoperability and reusability of data

To continuously expand 4Memory's emerging ecosystem of services while increasing the interoperability of services and reusability of data, TA3, with the other TAs, establishes a counselling service to support researchers and project leaders in applying for and carrying out projects and to advise them on selecting suitable services and data for their research questions. The service will consider standards identified or developed in TA2 and become a part of the 4Memory helpdesk organised by TA6-M4. M4 accomplishes three tasks (related problem stories: PS 1, 15, 17, 34, 45, 46, 48, 49, 56, 65, 67, 71, 75, 94):

TA3-M4-T1: Establish a technical counselling service: TA3 identifies our community's technical counselling needs and implements a corresponding service to provide an overview of DH tools and repositories, guidelines for selecting appropriate DH tools and repositories for specific research projects (e.g. tech papers as a series for the design of infrastructures for projects, developed with TA4-M4) and support for using services and standards. Consultancy also extends to the implementation of the policy framework for rights management (M3-T4).

TA3-M4-T2: Provide information materials: Conception of mediation formats for communicating the service portfolio and possible applications to relevant user groups. Design and implementation of information materials such as handouts, guidelines and best practices.

TA3-M4-T3: Networking: With regard to the services offered by 4Memory, the aim is to network, harmonise and coordinate with relevant actors at existing decentralised research data centres (FDZ) or research data management centres (FDM).

Contributing members: *Co-applicants*: FIZ, IEG, LABW, UT; *Participants*: HLA, LRZ, VdA, VZG, ZDV.

FIZ and LABW establish and coordinate the technical counselling service as part of the 4Memory helpdesk managed by TA6. Participants complement the service with specific consulting and support services related to long-term archiving (HLA, UT, VdA, ZDV), FAIR software (LRZ) and vocabulary management (VZG).

Measure 5: 4Memory Data Lab

The 4Memory Data Lab provides a technical infrastructure for the evaluation and deployment of innovative technologies and applications for annotation, feature detection and data analysis as well as text recognition. It will combine seamless access to relevant data sets and corpora via the Data Space with access to a high-performance hardware cluster. Suitable applications and technologies based on machine learning algorithms are developed or defined in the Method Lab (TA5-M2), including support for researchers on how to employ these emerging methodologies. The resulting annotations and deep indexing will flow back to the data-providing institutions via the virtual indexing layer and be available to third parties. In the first funding phase, the technical design is defined and implemented along practical requirements. Subsequently, the Data Lab will be put into operation. Starting point is the ML cluster of LABW and ZDV-T. In future, a distributed approach with further clusters and corresponding operators and institutions will be pursued. M5 involves three tasks (related problem stories include: PS 41, 64, 67, 68).

TA3-M5-T1: Establish Data Lab Board: Experts from FIZ, LABW, ZDV-T and HU form a board to coordinate and control the Data Lab's establishment and operation. It first elicits requirements for the ML cluster in coordination with TA4 and TA6. In later operation, it is responsible for adapting the framework conditions to evolving requirements. It also supports use of the Data Lab by providing advice and guidance.

TA3-M5-T2: Setting up and operating the Data Lab: An existing high-performance hardware environment jointly operated by LABW and ZDV-T serves as the Data Lab's technical basis. It consists of several machine-learning systems to be expanded by additional nodes across the funding period. Necessary computing power is provided by high-performance graphics cards (GPUs) and powerful CPU computing nodes. We use container-based virtualisation solutions such as Docker or Kubernetes to enable rapid deployment. Technical interfaces for data access to the Data Space are implemented.

TA3-M5-T3: Improve the retrievability of source contents via full-text recognition: TA3, with READ-COOP and SBB-PK, analyses available tools for optical character recognition, including handwritten, multilingual and non-western text resources. We provide best-of-breed solutions as part of the Data Lab, facilitating a steadily growing number of historically relevant sources available in full-text for text- and data-mining procedures and deep indexing via TEI. It will also improve the discoverability of digitised cultural assets in archives and libraries.

Contributing members: *Co-applicants*: FIZ, HU, LABW, MLU-HistData; *Participants*: SBB-PK, SLUB, ZDV-T

LABW and ZDV-T provide IT resources, FIZ expertise on deployment architecture, SBB-PK and SLUB OCR services. HU acts as liaison to TA5's Methods Innovation Lab and MLU-HistData to TA2.

Measure 6: Design and development of a virtual indexing layer

TA3 works with research projects, memory organisations and citizen scientists to design and implement a virtual indexing layer, combining memory institutions' indexing information (source descriptions) and deep indexing from projects (content annotation). The virtual layer enables the combined data's comprehensive search and retrieval, fostering reuse of valuable indexing information from projects. It will enable cooperative indexing or enrichment of existing indexing data that can be reintegrated into the original systems. We aim for easy transfer of annotations and metadata generated within the 4Memory Data Lab (M5). In the first funding period, TA3 develops and coordinates a concept with the community: implementation occurs in the second funding phase. M5 undertakes four tasks (related problem stories: PS 39, 54, 59, 73, 80, 85):

TA3-M6-T1: Collect and analyse requirements: LABW will extract requirements from problem stories collected during the application phase. In workshops with interested parties of the 4Memory community and with adjacent NFDI consortia, LABW will then collect further requirements for a virtual development level and, finally, analyse and aggregate them.

TA3-M6-T2: Conceptualise quality criteria and tools for specification and display: With TA1, LABW develops criteria for combining indexing information from research projects and catalogue information from memory institutions transparently and in ways comprehensible to researchers. Expertise from established citizen science projects will also be involved. Topics include handling uncertainties, hypotheses and errors and documenting the original informational context to enable source criticism even in a digital context.

TA3-M6-T3: Develop a system architecture related to the Knowledge Graph: LABW and FIZ will design a system architecture for the virtual indexing layer, including technical interfaces to the 4Memory KG and the 4Memory Data Lab.

TA3-M6-T4: Validate the implementation concept: LABW will discuss and validate the elaborated concepts in workshops with the community. Based on a prototypical implementation, LABW and FIZ will test the practicality of the concept together with a selected project.

Contributing members: *Co-applicants*: BSB, DM, FIZ, LABW; *Participants*: DNB, ExClu AM, HU-CMS, MPIWG, OFU, SBB-PK, SLUB, UE-FZG, VZG

FIZ and LABW establish and coordinate the expert group, organise workshops and manage the development of the virtual indexing layer. BSB, DM, DNB, ExClu AM, HU-CMS, MPIWG, OFU, SBB-PK, SLUB, UE-FZG and VZG contribute their broad and varied technical and conceptual expertise and support community involvement.

5.3.3 Milestones and Deliverables

Overview table on major milestones (MS) and deliverables (D) and the scheduled completion (project month).

Type	No	Title	Month
MS	1	Evaluation of APIs and viewer technologies completed	18
D	**I**	**Interface specification document**	**30**
MS	2a	Technical foundation for 4 Memory Knowledge Graph established	24
MS	2b	Knowledge Graph – API implemented	36
D	**II**	**First release of distributed editing environment for KG**	**48**
MS	3a	Criteria of an interoperable “Data Space” specified	24
MS	3b	Federated services for long-term archiving of research data available as key services	24
MS	3b	Key Services adapted to Data Space	36
D	**III**	**First release of discovery search service**	**42**
MS	4a	Network of relevant actors established	36
MS	4c	Information material created	40
D	**IV**	**Technical counselling service goes live**	**48**
MS	5a	Data Lab Board established	30
MS	5b	Full-text recognition/OCR services available as key services for Data Lab	36
MS	5c	Hardware environment is adjusted	42
D	**V**	**First release of Data Lab**	**48**
MS	6a	Requirements for the Virtual Indexing Layer collected	40
MS	6b	Quality criteria for combining indexing information defined	48
D	**VI**	**Concept document for virtual indexing layer**	**60**

5.3.4 Cooperation with other task areas and cross-area dependencies

A major task of TA3 is technically implementing the methods, concepts and contents developed in other TAs, making intra-consortium cooperation essential. TA3 will work with TA1 to derive FAIR criteria for historical research data, particularly to ensure the interoperability of 4Memory’s service portfolio. A key aim will be determining the data types and formats to be supported, which will define the requirements for the Data Space’s technical interfaces. The data model will be jointly developed with TA2, defining the community-specific, comprehensive authority data and vocabularies. The resulting ontology forms the conceptual basis for implementing the Knowledge Graph or the semantic search in the Data Space. TA3 will collaborate with TA2 on technologies and tools to extend the data basis (full-text recognition) and information extraction (entity recognition). TA4’s imparting of data/digital literacy is vital to TA3: 4Memory services will require

both digital knowledge and the skills of the historian. TA5's Methods Innovation Lab is complementary to the Data Lab. Both TAs promote using innovative digital services and foster a distinct data culture in historical disciplines. Finally, TA3 contributes its technical counselling service to the global helpdesk of TA6 and its expertise on sustainable operation models.

5.3.5 Quality Management and Risk Analysis

TA3's ambitious work programme entails risks that we will seek to mitigate. First, the spectrum of relevant services and offerings is technically and organisationally varied, with many involved parties, requiring substantial coordination of many participants and examination of a range of technologies and software. We address this by selecting key services to cover the community's basic needs provided by participants committed to work closely together. Second, technological progress ensures there will be great changes in needs and solutions during the project. We will work closely with service providers and the community via our working groups to continuously adjust the work programme as changes arise. Third, the funds applied for only allow a basic achievement of objectives, especially for M3 and M6, and must be accompanied by third-party funded projects. Acquiring these funds cannot be taken for granted. In M4, therefore, we will support third-party funding applications to improve their chances through advice, coordination and content-based participation. Fourth, heterogeneous data quality may impact development of a knowledge graph, which requires meaningful data. We will compensate through entity recognition and disambiguation algorithms to semi-automatically enrich metadata with needed semantic annotations. Fifth, we cannot meet all our community's needs, particularly regarding continued operation of services from expired projects and the rescue of orphaned data. Expectations must be mitigated and addressed in the second funding period.

Page intentionally left blank

5.4 Task Area 4: Data Literacy

5.4.1 Description of the Task Area

Building on the needs described by our community, TA4 develops specialised methodological and curricular innovation to enhance data literacy (DL) in historically oriented research and teaching. This is a core prerequisite for the realisation of the much-needed cultural change in our profession (see TA5). TA4 and TA5 thus build on one another. They differ, however, with regard to their overarching goals and the measures taken to reach them. TA4 promotes data literacy by focusing on competence development and skills acquisition. It aims at integrating (i.e. “mainstreaming”) the teaching of data literacy at all levels of history education (regular BA, MA and PhD programs as well as area history courses, e.g. in African, Asian, European and East European, Latin American and Russian history). TA4 thus counterbalances current tendencies to compartmentalise data literacy training by creating ever more specialised Digital History programmes. TA4 also aims to normalise global history approaches by systematically integrating chronological and geographical perspectives in our expert discussions to innovate the “historical method of source criticism” under the digital condition.

DL and, specifically, data criticism, understood as “the ability to collect, manage, evaluate, and apply data in a critical manner” (Ridsdale et al. 2015, 2), are essential skills for nearly all sectors and disciplines of the global knowledge-based economy. However, DL is of specific relevance for historically oriented scholarship since it distinctively affects the discipline’s methodological self-conception – the “historical method of source criticism”. Source criticism is the process of evaluating the qualities of an information source, such as its validity, reliability and relevance to the subject under investigation. Sound historical scholarship depends on a rigorous, careful application of the “critical historical method” as the basis for historical hermeneutics and the intersubjective validation of historical knowledge production. Historical source criticism takes into account that each individual historical source has a specific origin and tradition history. The reconstruction and critical evaluation of this history is a central element of source contextualisation, which always stands at the beginning of the historical research process. Digitalisation and datafication of historical source material adds additional layers to the history of the origin and tradition of a historical source. Data modelling further complicates historical source criticism in that the critical evaluation of a data model asks for skills that historians usually do not acquire during their scholarly education, such as data modelling standards, coding and script languages and information technologies. Therefore, the method of “historical source criticism” must be expanded by including skills in data criticism. Historians must be able to critically analyse the data models, metadata, infrastructures, IT systems, interfaces, and algorithms that are used in the process of creating historical data. They need to be able to evaluate them with regard to their implications for information retrieval and their impact on research results.

While the growing number of digitised and born-digital sources representing the history and traditions of diverse societies poses many challenges to the historical method, raising questions about culturally distinct as well as time-specific ontologies and epistemologies in the digital age, they offer innovative methodological opportunities going beyond the classic hermeneutic approach of the historical method. Digital instruments and tools such as digitally supported spatial, textual and network analyses; simulation, sound and image analyses; and methods already established in other disciplines, such as text mining and topic modelling promote innovation in many historical fields and subdisciplines. The dynamic of the innovation processes in historical scholarship engendered by the digital turn depends, however, on historians' abilities to use digital sources and apply digital tools. To fast-track innovation, TA4 develops DL training material and services that enable historians to evaluate digital sources (including algorithms, interfaces and infrastructures) and to reuse historical research results produced with digital methodologies and based on digital source material according to the FAIR and CARE principles. This includes acknowledging cultural, temporal and linguistic specificities and taking into account issues of data ownership, sovereignty, ethics and rights. As a contribution to capacity building, TA4 establishes counselling services that offer easy access to new digital instruments and tools and their application. By reaching out to the broad history community, TA4 contributes to disseminating and mainstreaming DL competencies and skills at all levels of digital training: programmes to "teach the teachers", doctoral and post-doctoral education, BA and MA programmes, including teacher training programmes and area history programs. Our measures contribute to the LINKAGE objectives of ***I****ntegrating historical source criticism into data services*, building a ***N****etwork of historically oriented research communities* and furthering ***E****ducation and citizen participation*.

The rich, growing reservoir of historical data and the increasing number of digital tools and instruments multiplies the roles that historically oriented scholars at universities, archives, libraries and infrastructure institutions play in different stages of the historical research process. TA4, therefore, addresses the skills necessary to master the whole research cycle, from exploring an idea, finding background information, locating current research and identifying relevant historical sources to evaluating, analysing, interpreting and citing those sources (see the DIAMANT Competence Matrix in Lemaire et al. 2020, 29–40). However, despite the growing number of digital sources, the larger part of the historical cultural heritage will remain accessible only in an analogue format, and historians will continue to work with analogue archival materials and library holdings. Thus, with regard to our LINKAGE goals, TA4 also offers services to help historians ***A****dvance the analogue / digital interface of historical source material and data*.

5.4.2 Measures

Measure 1: Data Literacy Training Lab

M1 links experts in the field of PhD and doctoral education, historical didactics, teacher-training institutions and secondary schools who collaborate in promoting curricular innovation and conceptual progress in teaching DL and RDM competencies. By transferring expert knowledge into academic study programs, school curricula and the curricular norms and standards defined by state ministries, M1 promotes a target-group-specific mainstreaming of data literacy education and training for secondary school teachers, BA and MA students, PhD and postdoc researchers. This includes topics such as data organisation and manipulation, data conversion, metadata, data curation, reuse and preservation (in cooperation with TA1, TA2, TA3) as well as algorithm, interface and infrastructure criticism. M1 develops innovative DL-training modules and material (T2), proposes a curricular model to mainstream DL-training in BA and MA study programmes (T3), and establishes a DL-accreditation service (T4) for quality management and advocacy with responsible political actors and agencies.

TA4-M1-T1: DL-competence profile: Identifying necessary competencies and skills for historically oriented humanities (in cooperation with information scientists); mapping and evaluating existing training material; establishing a DL training catalogue comprising didactic and content recommendations.

TA4-M1-T2: DL-training modules and material: Developing and implementing target-specific modules and training material addressing history students in secondary schools, BA and MA programmes (including teacher training programmes), and PhD and postdoctoral education.

TA4-M1-T3: Mainstreaming DL-competence education: Developing a model curriculum for BA and MA study programmes, which systematically integrates the teaching of DL-competencies into basic and advanced courses and seminars; implementing pilot programmes at cooperating institutions; evaluating and disseminating the model curriculum.

TA4-M1-T4: DL-accreditation service, quality management and advocacy: Developing quality criteria for DL training material and guidelines; establishing a DL accreditation service; disseminating didactic and content recommendations; initiating advocacy with academic and teacher-training institutions, secondary schools and state ministries.

Participants FU-CEDIS and UzK-ZfL coordinate the tasks, prepare the training-material and establish the accreditation service. They are also responsible for the creation of an advocacy and outreach strategy addressing the politically responsible public agencies. They are supported by *co-applicants* UT, HU, h_da, HI, IEG and *participants* ArchivSch, DHI Moskau, DHI Rom, DHI Washington, HDU, UL, BUW-DH. Participants from the field of history didactics (FU-Didaktik, KGD, HDU-HSE) monitor and evaluate the pilots. Research results will be published in hybrid

formats to reach the diverse group of stakeholders, from teacher training to university education to state ministries. Co-applicant VHD and participants MV, MGH, VdA support the advocacy policies of M1.

Contributing members: *Co-applicants*: UT, h_da, HI, HU, IEG; *participants*: ArchivSch, BUW-DH, DHI Moskau, DHI Rom, DHI Washington, FU-Didaktik, HDU, KGD, UL, UzK-ZfL.

Measure 2: Historical Data Criticism and RDM

M2 establishes an expert group for historical data criticism and research data management, focusing on time and area-specific research-based data collection, production, management and evaluation as well as the analysis and interpretation of historical data. M2 identifies challenges in using, presenting, managing and storing historical data on specific historical periods, from antiquity to contemporary history, and on different cultural and knowledge contexts, developing guidelines for critically handling and (re)using data-driven research results. To reach out to our community of interest and to secure the participation of national and international experts, M2 establishes a blog (via hypotheses.org) on “The historical method under digital conditions”. It creates time- and area-specific syllabi as best practices demonstrating solutions to the challenges of teaching historical data criticism and sustainable RDM. It develops guidelines for the historical contextualisation of unstructured and/or heterogeneous historical data and innovative training material for using digitised sources in the history classroom. These deliverables feed into the publication of a textbook on historical data criticism and RDM covering a broad spectrum of historical topics and addressing politically relevant actors such as accreditation agencies, state ministries, and educational boards. M2 undertakes the following tasks:

TA4-M2-T1: Guidelines for historical data criticism: Organising an interdisciplinary international conference on research in historical data criticism, digital hermeneutics and the critical handling and (re)use of data-driven research results; supporting interdisciplinary communication and exchange on data criticism and sustainable research data management; internationalising the expert group’s expertise by including international experts.

TA4-M2-T2: Blog “The historical method under digital conditions”: Promoting national and international exchange about further developing the historical method under digital research conditions; disseminating recommendations and research results.

TA4-M2-T3: Syllabi and training material: Developing innovative time- and area-specific syllabi and training material as best practices to demonstrate solutions to the challenge of teaching history with digital source material and to promote the mainstreaming of digital methods education in regular history programmes, including area history courses.

TA4-M2-T4: Historical contextualisation of data models: Developing time-specific recommendations and guidelines for the digital contextualisation of unstructured and/or

heterogeneous historical data as a central element of historical data criticism, sustainable RDM and the further development of the historical method; creating and publishing textbook knowledge.

Co-applicant UT and *participant* DHI Washington coordinate the tasks. They are supported by co-applicants h-da, HI, HU and IEG. *Participants* DHI Rom, DHI London, HDU, UL, ExClu AM, UT assure that the expert group covers both the *longue durée* and the area history perspective of historical research. They contribute to curriculum development by piloting DL-sensitive module descriptions and establishing innovative syllabi for courses in basic and advanced training to mainstream the teaching of data literacy in history education.

Contributing members: *Co-applicants*: h_da, IEG, HI, HU, UT; *participants*: ExClu AM, UT-TRANSMARE, UT-SEAL, DHI Rom, DHI London, DHI Washington, HDU, UL, UT.

Measure 3: Digital area histories

M3 addresses technical, legal and ethical dimensions of historical data criticism and RDM in area histories (African, Asian, European, East-European, North American, Latin American and Russian History). Digital area histories face distinct challenges because creating and using research data must be coordinated with the national and regional institutional, technical and legal contexts in which these data were primarily created, collected, made accessible and managed. Therefore, RDM in area histories need to consider issues of technical fit between different traditions of data collection and storage as well as indexing through metadata standards. It must also account for divergent legal regimes of data access, research data use and good scholarly practice. Finally, yet importantly, Europe's colonial history and treatment of its colonial historical heritage as well as persistent global asymmetries of academic research opportunities, pose specific issues of research ethics in area histories. M3's central objective is twofold: (1) to map and monitor such differences and make available reliable and relevant information on the international development of research data management for the German research and education landscape; (2) to initiate and document the negotiation of cooperative procedures for using data from other national and regional contexts and to thus make this data accessible for German scholars in a "fair" and "careful" manner. This leads to the following tasks:

TA4-M3-T1: Area-specific needs assessment: Initiating a transdisciplinary community network to map and assess the specificities of border-crossing research approaches using archives and data from different countries and often teaming-up or informally cooperating with scholars from other countries; developing and publishing area-specific recommendations for digital source and data criticism and sustainable RDM.

TA4-M3-T2: Technical regulations and RDM standards: Establishing an information service about technical trends and standards in other countries regarding access, storage, and digitisation of data for historical research; developing RDM plans for area histories that (a)

address the diversity of technical regulations and data-quality standards in different countries and (b) are based on cooperation with scholars from these countries; establishing DL Labs for junior and senior scholars focusing on area-specific RDM.

TA4-M3-T3: International legal frameworks and regimes: Building on M3's area-studies thematic focus: providing scholars in German institutions with appropriate, up-to-date information on international legal trends, features and regulations determining access, collection, proceeding, (re-)use and storage of historical data from non-European archives and information infrastructures; formulating recommendations for historical research under (sometimes radically) different legal frameworks; collaborating with TA5 on promoting an international data culture that adheres to NFDI rules but enables fruitful compromises with non-German data cultures.

TA4-M3-T4: Ethical challenges of digital area histories: Identifying specific ethical aspects of research on transborder topics and the use of archives and data across borders; publishing a white paper on questions such as: who "owns" data, who determines which questions are raised and answered with data, who cares about the priorities in data protection and digitisation of archives, and how to build cross-border communities able to use data for jointly defined purposes; training junior and senior scholars in ethically approaching data cultures in various world regions.

Participant UL coordinates the tasks. UL is supported by *co-applicants* FU, IEG, HDU, HI, HU and UT. In addition participants CERES, DHI Moskau, DIJ Tokyo, ExClu AM, FU-LAI, FU-UB, MWF Delhi, GU-UB, IAI, IGDJ, OIB contribute (a) specific area and transcultural expertise, (b) know-how in dealing with multilingualism, non-Latin scripts and characters, (c) familiarity with technical and legal challenges of cross-cultural historical research and (d) a network of international scholars to contribute to the development of area-specific RDM-plans.

Contributing members: *Co-applicants*: IEG, HI, HU, UT; *participants*: CERES, DHI Moskau, DIJ Tokyo, ExClu AM, MWF Delhi, FU-LAI, FU-UB, GU-UB, HDU, IAI, IGDJ, OIB, UL.

Measure 4: Digital research counselling service

To empower historical researchers, instructors and trainees to conduct digitally driven historical research, M4 will establish a counselling service for digital methods, tools, instruments, research techniques and forms of collaborative research data management. The counselling service will connect researchers (a) with experts working on and with digital tools and (b) with professionals at libraries, archives and infrastructure partners. Based on a needs and qualification assessment, M4 will develop virtual training materials (webinars) and pilot regional training centres and contact offices to serve as the basis for establishing a virtual DL counselling service and helpdesk. Webinars and onsite peer-to-peer consultation support the training of future trainers and users. The training courses and training materials as well as information about digital tools, instruments, research techniques, local and regional contact offices and the helpdesk will be published on the

4Memory portal (DL virtual hub). The counselling service will become part of 4Memory's service portfolio.

TA4-M4-T1: Community involvement and needs assessment: Securing the continuous exchange between community members working on and with digital tools by setting up an organisational structure to offer workshops and coordinate road shows on DL tools, instruments, and research techniques over the whole funding period (in cooperation with TA5); assessing the needs and qualifications of the community.

TA4-M4-T2: Webinars for most-needed digital methods, tools, and research techniques: Developing virtual training courses (webinars) on the use and implementation of the most-needed digital tools, instruments, digital research techniques and RDM.

TA4-M4-T3: Regional training centres: Piloting regional training centres and contact offices to "train the trainers" and to offer onsite consultation for possible users.

TA4-M4-T4: Virtual counselling hub: Establishing a virtual hub for training courses and training material for DL services with information about local and regional contact offices. Developing piloted training centres (TA4-M4-T3) into a digital tools counselling service and integrating the service into 4Memory's service structure.

Co-applicants UT-SeS and h_da coordinate all tasks, supported by *co-applicants* MLU-HistData and IEG. *Participants* CERES, DHI Washington, ExClu AM, FB Gth, FU, GBA, MWF DELHI, UHH, UL, UzK-DCH, UzK-ZfL contribute their expertise.

Contributing members: *Co-applicants*: h_da, MLU-HistData, IEG, UT-SeS; *participants*: CERES, DHI Washington, ExClu AM, FB Gth, MWF Delhi, FU, GBA, UHH, UL, UzK-DCH, UzK-ZfL.

5.4.3 Milestones and Deliverables

Overview major milestones (MS) and deliverables (D) and the scheduled completion (project month).

Type	No	Title	Month
MS	1a	DL-training catalogue established	12
MS	1b	Target-specific modules and training material disseminated	48
MS	1c	Model curriculum developed	24
D	**I**	**Pilot programmes implemented**	**54**
D	**II**	**Governance model DL-accreditation service established**	**48**
MS	2a	International Conference "Historical Data Criticism" convened	12
MS	2b	Blog "The historical method under digital conditions" established	12
MS	2c	Time-specific syllabi developed	24
D	**III**	**Textbook "Historical Data Criticism and Sustainable RDM" submitted to peer-review**	**48**
MS	3a	Area-specific recommendations for data criticism and RDM published	48
MS	3b	Technical information service for Area Histories established	36
MS	3c	Legal recommendations for Area Histories published	24
D	**IV**	**White Paper on ethical dimensions of digital area histories published**	**48**
MS	4a	Needs assessment completed	12
MS	4b	Virtual training courses developed	48
D	**V**	**Regional training centres established**	**24**
D	**VI**	**Virtual counselling hub launched**	**36**

5.4.4 Cooperation with other task areas and cross-area dependencies

Data literacy training is essential to fast-tracking innovation in historical research. TA4's measures and tasks will promote the digital transformation of historical research and teaching. TA4 thus closely cooperates with TA5 and with TA1 and TA2. Whereas TA5 aims at cultural change, TA4 targets capacity building. TA1 and TA2 focus on critically reflecting upon the relationship between data, authority data, metadata and historical knowledge; TA4 addresses the competencies, techniques, training materials and guidelines historians need to immerse themselves in using digital sources for researching and analysing past events, processes and developments. Our measures cover aspects of DL related to the quality and specificities of the broad spectrum of heterogeneous and often unstructured source material and data domains historians use in their research: texts and documents, material objects, oral histories and other sound-sources, images and other kinds of visual representations of the past, as well as born-digital sources, including sources of non-European world regions.

Besides contributing to the LINKAGE objectives, TA4's deliverables will add to achieving the tasks and objectives of the newly founded NFDI sections on ethical and legal aspects and on education and training. By highlighting DL as a key issue and addressing its core challenges in a distinct TA, 4Memory will establish instruments and services especially relevant to the consortia Text+, NFDI4Culture and NFDI4Objects, with whom TA4 has already established close contact during the preparation process. Regular meetings of the responsible spokespersons will ensure continuous communication and exchange on DL to establish inter-consortial service and quality-management structures as well as a discipline-independent DL competence profile.

5.4.5 Quality Management and Risk Analysis

The co-spokespersons of TA4 are responsible for its quality management (QM), supported by a Task Area Steering Committee composed of the responsible coordinators of the four measures. TA4 follows an object-oriented QM model. Starting from cooperatively defined quality standards for each measure, we will formulate quality indicators to evaluate and control the processes and their outcomes. Since TA4 aims to establish several services and products – such as teaching material, syllabi, model curricula, webinars, etc. – its quality management focuses on product and service quality and monitors the means and processes to achieve it. An important element of TA4's quality management is setting up a permanent feedback mechanism involving our target groups. Counselling services and regional contact offices facilitate the direct exchange with the users of our services and products and connect them with the service providers, giving them both outreach and internal feedback functions. The co-spokespersons of TA4 cooperate closely with the quality management structure established in TA6 and are members of 4Memory's Quality Management Board.

We have identified four potential risks facing TA4's goals and objectives: (1) the limited time budget of scholars and history teachers, which may prohibit the timely implementation of new syllabi and training material in daily teaching and research routines; (2) insufficient financial resources for establishing the necessary services for capacity building at research and teaching institutions; (3) duration of the examination and approval processes of teaching and learning materials and study regulations by the responsible authorities or university administrations; (4) lack of interest by scholars in other world regions to cooperate with us on ethical issues of digital Area Histories. Given the opportunities identified in our analysis, we are convinced we can manage such threats. With the help of our DL Training Lab, our Regional Training Centres and our Virtual Counselling Hub, we will counterbalance the identified weaknesses and threats by proactively addressing the growing scholarly, institutional and political interest in establishing RDM services, using digital methodologies and establishing digital research and teaching material. The inclusion of well-established German centres for Area Histories as well as the German Historical Institutes in non-European world regions as participants will offset possible problems in accessing scholars from other world regions.

5.5 Task Area 5: Data Culture

5.5.1 Description of the Task Area

TA5 aims to move our community towards a data culture that integrates digitalisation in historical fields and responds to its challenges proactively, innovatively and responsibly. To reach this goal, TA5 cooperates closely with TA4: while TA4 imparts data skills and competencies to students, researchers, employees of memory institutions and others working with RDM, TA5 promotes community acceptance of research data and the changes it brings and raises awareness of the requirements and consequences of RDM among those who work with it and those who set the standards of professional status and advancement. This is essential: digitalisation means not only technology but also deep epistemological and social change in our disciplines. Integrating 4Memory's infrastructure into researchers' everyday lives means fully understanding this transformation: our community needs policies and guidelines for a data culture that values research data, facilitates and promotes data sharing and (re)use, and provides data security and

reliability. This culture must accord with the FAIR and CARE principles, ensuring data is used not only in a scholarly but also ethical way.

Broadly defined, “data culture” refers to integrating data and data-driven methods and practices into the historical disciplines’ professional culture: generating a wide-ranging discussion of how they change the production of historical knowledge and our disciplines more generally and increasing their understanding, acceptance and use. More narrowly, the term encompasses day-to-day social practices and responsibilities in dealing with research data, including legal and ethical issues, questions of sustainability and, most importantly, the recognition of data as scholarly achievements. We address both facets of “data culture”.

TA5’s work programme must account for the particular status of RDM in historically oriented research in Germany, where working with research data and applying data-driven methods to analyse them is still in its infancy. Knowledge about the concept, content or potential of research data in historically oriented disciplines has been scarce and unevenly distributed. There are reservations as to the impact of data-driven practices and methods on the production and provision of historical knowledge, which often stem from limited familiarity with these methods. Working with digital data and advancing technology leads to new ethical and legal challenges for researchers and society as a whole in dealing with our historical heritage, the full (and future) implications of which are not yet even understood. The creation and publication of digital research data, not least due to the lack of peer review, has yet to be recognised as an original scientific contribution, with considerable consequences for the career paths of young and innovative academics. Similar issues apply to new forms of collaborative research and authorship required in digital research.

Building on these observations, TA5 will enable **a discourse on change regarding epistemologies, knowledge orders and disciplinary contexts**, thus improving our understanding of digitalisation in our discipline (M1); evaluate, explain, and improve **access to current and future methods**, fostering digital collaboration and innovation in our field (M2); identify **legal and ethical challenges** associated with new methods and technologies and provide relevant guidance, increasing confidence and awareness in our community (M3); discuss and implement **new publishing and review practices** at the centre of our discourse, integrating research data into our disciplines’ proven, established quality management and reputation systems (M4); and provide the **infrastructure to enable networking, exchange, and communication** regarding data culture in our community and beyond (M5).

Creating a robust and adaptable cultural change means accounting both for current realities in our community and for future developments in an integrated perspective. TA5 places communication and exchange with our whole community – whether in research, memory institutions or public history – at its core (M1-T2, M3-T3, M5-T2) while also promoting pioneers in

the movement toward datafication (M1-T4, M2-T4, M5-T3). We will showcase examples of digital excellence and propagate solutions to obstacles in data-driven research. We must also take reservations, uncertainty and criticism seriously to develop a critically-minded, innovative and sustainable digital culture. Our deep embeddedness in our large, heterogeneous community will enable us to integrate it in developing our programme, which, where possible, will be driven by the community itself.

Shaping cultural change, we ***L****ink research, memory institutions and infrastructures*, build a ***N****etwork of historically oriented research communities* around a common data culture and ***G****enerate (professional) standards for historical research data and sustainability*.

5.5.2 Measures

Measure 1: Epistemologies, knowledge orders, disciplinary contexts

We must encourage reflection on digitalisation's consequences: changing epistemological foundations, the permeability and potential redrawing of disciplinary boundaries, and the role of citizen sciences in producing historical knowledge. Future research must account for the impact of the digital encoding of knowledge systems, e.g. cultural bias and constraints on long-term disciplinary development. (See problem stories: PS 80, 91, 92.) Addressing these questions, M1 provides a basis for other measures. The goals are: bringing together experts and gatekeepers from across the community, but also representatives from neighbouring consortia such as 4Objects, to further a critical discourse on digital epistemologies and disseminate their findings to the wider community; engaging in a vibrant exchange with the community; enabling a mutual integration between our discourse and those of the relevant international communities; and helping early-career researchers join the debate. Specific topics will be determined by the expert group and may include: the datafication of historical sources, digitality and epistemology; AI and historical knowledge production; the impact of knowledge graphs and authority files on the knowledge order; born-digital sources and the loss of authenticity; and new disciplinary boundaries and professional profiles.

TA5-M1-T1: Expert group on the consequences of datafication: Regular meetings of the group and subgroups dealing with key problems of digital change in the historically oriented disciplines, with the involvement of young scholars (T4) and including planning and coordinating for T2, T3 and T5. Support for the biennial VHD-DGW conferences on digital history to discuss epistemological changes with experts on a larger scale.

TA5-M1-T2: Roundtables with professional historians' associations: Integrating sessions and roundtable discussions at the conferences of professional societies and associations and fostering corresponding discourses in their journals, blogs, etc. to directly engage with the broader

community, involving them in the discussion so as to hear their concerns, reservations and questions, which are then taken up in T1.

TA5-M1-T3: International workshop series on research data and epistemologies: Co-organisation with TA4 of a workshop series at the DHI-Paris, DHI-Washington and MWF Delhi in exchange with the respective national and regional research communities.

TA5-M1-T4: Support of early career researchers: Supporting projects by early career researchers through a stipend to fund travel grants to attend expert-group meetings and conferences and to publish their findings (e.g. for language editing, OA fees).

TA5-M1-T5: Hybrid book publication: Publication of a collection of essays on the new digital foundations of the field of history, digitalisation and its effects on epistemology, knowledge order and disciplinary culture, summarising the results of M1. Through a hybrid publication, the findings shall be made accessible to the entire community, including those who are rather hesitant to consult digital publications.

Contributing members: *Co-applicants:* h_da, HI, HU, LABW, UT, VHD; *Participants:* AHiG, BerHis, BUW-DH, BUW-IGP, FU-CeDiS, DHI Moskau, DHI Rom, DHI London, DHI Paris, DIJ Tokyo, ExClu AM, MWF Delhi, FZH, DHI Washington, HU-CMS, IRS, MG, MV, RHohls/HU, SBB-PK, UBI, UE-GG, UJ, UL, UP-DH, VHD-DGW, ZZF

HU leads this measure in close coordination and with the support of VHD and provides the organisational framework and coordination. Representatives of institutions from across the consortium and from German-speaking countries abroad contribute expertise and experience to the expert group and participate in the intended book publication. Several German institutes abroad cooperate with additional workshops and VHD-DGW contributes with its biennial “Digital History” conference to facilitate discussion and opinion formation.

Measure 2: Methods Innovation Lab

Researchers will only create and provide research data if they know how to use them and are sure that others will be able to use them effectively. Relevant methods must be made known and accessible and anchored in practice. We first gain an overview of the state of the use of data and methods in historical studies, track community developments and trace 4Memory’s impact and the need for further action. We evaluate established and innovative methods, demonstrate their utility and make them available to researchers with the necessary data skills (TA4), while creating incentives for others to acquire them. By drawing on the consortium’s infrastructure and data (with TA3), we create use case demonstrators. By organising hackathons, impulses will be gathered from a broader community, highlighting the public relevance of the critical function of research data and the digital transformation of historical studies. The information and competencies thus gained will be proactively made available to DH centres and coordinators at universities, research

and memory institutions and, through them, anchored in practice; in return, their experiences will be reintegrated in Task 2. Finally, we shift the focus from methods to research questions to incorporate those methods in current historical research, strengthen the link and dialogue between analogue and digital historians and showcase new methods and practices (such as co-authorship) in traditional publications. In doing so, we aim to boost disciplinary innovation processes while ensuring domain-specific accessibility to a variety of algorithms and methods to prevent monopolisation, which is essential for free and open research and use of research data.

TA5-M2-T1: Survey of digital methods and their application in historical research: Lab staff produce and publish a survey on the current use of research data and digital methods in historically oriented disciplines after the 1st and 4th year.

TA5-M2-T2: Exploring and promoting new methods for data driven research: Together with the community and in close exchange with consortia from related fields (e.g. 4Objects, in their Task Area 6), Lab staff identify and evaluate established and innovative methods, make them known via exemplary data stories and, accounting for domain-specific needs, provide tutorials and code examples as Jupyter notebooks on the Clio Guides Platform (M5-T1) in close cooperation with the 4Memory Data Lab (TA3).

TA5-M2-T3: History hackathon: At least two hackathons based on digitised and digital-born research data relevant to public history will be organised, inviting participants also from other fields and their ability to innovate and raising public awareness of the current changes in historical scholarship and its relevance to society.

TA5-M2-T4: Establishing exchange with DH centres and coordinators: Setting up communication structures, such as newsletters, to inform DH centres and coordinators about the 4Memory Method Lab, providing advice, enabling exchange on digital methods and integrating their solutions in our portfolio of data stories and code examples (T2).

TA5-M2-T5: Promoting collaborative research: community projects on specific case studies: Call for projects to identify 3 or 4 research questions suitable for using research data and digital methods, matchmaking with experts and support for corresponding collaborations through travel funds, research software engineering and publication of the results.

Contributing members: *Co-applicants:* h_da, HI, HU, LABW, UT, VHD; *Participants:* AHiG, BBF, BerHis, BUW-DH, BUW-IGP, FB-Gth, FU-CeDiS, DHI Moskau, DHI Rom, DHI London, DHI Paris, DIJ Tokyo, DSM, ExClu AM, GEI, DHI Washington, HU-CMS, IRS, MG, MV, SBB-PK, UBI, UE-FZG, UE-GG, UL, UP-DH, VHD-DGW, ZZF

HU provides the organisational framework and coordination of this measure, in dialogue with institutions and scholars from across the consortium, who provide data and data stories, contribute code and tutorials and act as partners for collaborative projects.

Measure 3: Ethical and legal implications

Responsibly using sources from contexts of injustice (such as dictatorships, colonial rule or genocide) has long been central to the historical study of entire eras, regions and themes. But the changes that come with the digital representation of these sources, born-digital sources and new technologies – such as the decontextualisation of data or their easier availability, duplicability and mutability – pose new challenges for our field and for society. We need a comprehensive discourse on the ethical principles and legal challenges arising from Machine Learning, Big Data and the emergence of AI (such as synthetic data, deep fakes or bias); on public responsibility for open access research data and its reuse by third parties with different agendas (such as private companies or government agencies); on data authenticity and falsification; on protecting personal data and privacy; on data security; on copyright issues and the legal framework for preserving and providing digital sources for future research (to avoid the loss of the history of the digital present); on issues of scientific freedom such as user tracking or free access to data and open software as well as the ecological footprints of our research data. Many in our community are unsure about these issues: moving towards a data culture thus requires a better understanding of and competence in addressing them. We will gather a range of expertise and also explore the legal and ethical challenges arising from technological developments (T1) to provide guidance to the community and the public and to highlight where steps must be taken (T2). The often abstract and complex nature of such issues has limited the community's interest in addressing them. Thus, they need to be communicated as comprehensibly as possible (T3). Building on this, the community needs basic information (T4). Finally, since historical data and their handling are also societal issues, this discussion must extend beyond narrower academic circles into the public (T5).

TA5-M3-T1: Expert working sessions on future problems and challenges: Group of experts consisting of members of the Advisory Board (TA6), representatives of community stakeholders and other specialists meet regularly to discuss emerging ethical and legal issues in handling historical data (particularly those arising from new technologies) and coordinate actions and initiatives (T2, T3, T4, T5).

TA5-M3-T2: White papers on future problems and challenges: The expert group publishes white papers on selected legal and ethical challenges arising from new technical methods with recommendations for action addressed to our professional societies as well as to policymakers.

TA5-M3-T3: Illustrating complexity – multimedia stories with exemplary projects: Identification of suitable projects with complex but typical legal and ethical challenges and development and publication of the most suitable formats (blog post, multimedia story, podcast) for their exemplary communication.

TA5-M3-T4: FAQs on ethical and legal matters: Collect and publish fundamental questions about and answers to legal and ethical issues in the Clio Guides (M5), building on existing, though dispersed, information and collaboration with other NFDI consortia under advice from our community's expert group (M1).

TA5-M3-T5: Public lecture series: Organisation of two public lectures per year in cooperation with experts on data ethics and data law outside the field of historical science.

Contributing members: *Co-applicants:* BSB (FID History), FIZ, h_da, HU, MLU-HistData, UT, VHD; *Participants:* AK-PF, ArchivSCH, BBF, BerHis, FU-CeDiS, DHI Moskau, DHI Paris, DIJ Tokyo, FZH, GEI, DHI Washington, RHohls/HU, MWF Delhi, IFZ, IRS, UBI, ExClu AM, UE-FZG, UL, VHD-DGW, ZZF

The VHD leads M3 in collaboration with HU, providing the organisational framework and coordination. It works with experts across the consortium (ZZF, FU-CeDiS, BBF, MWF Delhi etc.), in legal matters especially with FIZ and HU. For the public lecture series we will seek collaboration with established experts in the field such as the *Rat für Digitale Ökologie* and *Re:publica* to bring our agenda to the public.

Measure 4: Publication and review culture with regard to research data and data-driven research

Researchers' status and visibility depend on peer review, which ultimately defines "good" and "relevant" research. Increasing the acceptance of data-driven research in our field requires incorporating a review system for research data and data-driven research into practices and institutions established over centuries. Digital transformation and datafication are leading to even greater changes in our publication culture, as established genres expand via data publications, data papers or computational publications and digital publications of all kinds. These shifts entail new forms of authorship, publication technologies and citation methods. A new data culture requires a new publication culture and – especially for innovative young researchers – integrating research data into our established reputation system. M4 identifies challenges and introduces solutions (T1), testing and implementing them via our own publications, exemplarily anchoring them in the publication culture (T5). M4 focuses on integrating research data and data-driven research into the reputation system, requiring new assessment criteria that combine established, analogue standards of quality assessment with technical, data-driven methodological competencies (T2). Changing review processes and their technological underpinnings must also be explored (T3). Relevant experience of editors and journals (such as the *Journal of Digital History*, H-Net etc.) will be sought out where possible. Above all, changes must be visibly implemented in the key review publications, as is done here using the online platform H-Soz-Kult, published by Clio-online e.V., with its over 30,000 subscribers, over one million monthly

pageviews by natural persons, and more than 1,000 reviews per year covering all fields of historically oriented humanities (T4).

TA5-M4-T1: Expert network "Publication and Review Culture": A network of experts and stakeholders meets regularly to identify and discuss best practices for data publications and enhanced publications, new forms of co-authorship and citation and their integration into bibliographic reference systems and peer review; they are published on ClioGuides (M5). May include advising professional societies and producing white papers on specific topics.

TA5-M4-T2: Establishing evaluation criteria for the review of research data: Development and publication of a criteria catalogue for the (peer) review of research data and research data based publications by the expert network in close cooperation with TA1 and the involvement of the community, its testing and iterative improvement.

TA5-M4-T3: Development of standard editorial workflows for reviewing research data: Analysis of requirements for editorial structures, competences, technical features, review formats and incentive systems to adopt the standard workflow to integrate reviewing research data and research data-based publications into the existing reviewing system.

TA5-M4-T4: Development of an exemplary review module @ H-Soz-Kult: Agile development and implementation of a module for reviews of research data and research data-based publications on the central community-driven review platform H-Soz-Kult and integration, improvement and showcasing the results of T2 and T3.

TA5-M4-T5: 4Memory open access publication series: Setting up an open access publishing environment, establishing editorial processes and providing editorial support for publications emerging from 4Memory with a series for data papers as showcases.

Contributing members: *Co-applicants:* BSB, DM, h_da, HI, HU, MLU-HistData, UT, VHD; *Participants:* Clio-online, DHI Paris, DHI Rom, DSM, GEI, DHI Washington, HU-CMS, IDE, RHohls/HU, UE-FZG, UE-GG, UBI, UL, UL-AG, UT-SeS, VHD-DGW, ZZF

Clio-online coordinates M4, supervised by VHD and HU, and, with H-Soz-Kult, provides the infrastructure, existing expert networks and editorial framework for implementing a data-review segment. UT, h_da, UT-SeS, IDE, HI, UE-GG and others contribute to the expert groups, ZZF and GEI also provide work time for the planned editorial group on research data.

Measure 5: Creating impact, shaping change: enabling critical discourse in our community

What practical means will shift the historically oriented humanities towards a data culture, increase the acceptance of research data and support the next generation of digital researchers in working with research data? Other measures enable reflection, methodological innovation, discussion of ethical and legal issues and increase recognition for data-driven research: M5 aims at implementing these efforts among different groups, especially early-career researchers. It

provides an infrastructure for informing, discussing, networking and collaborating, thus anchoring needed changes in our community. It includes creating a communicative hub on data culture (T1); providing space and time for master's students, early career researchers and practitioners from academia, memory institutions and public history and citizen science projects to learn about digital transformation, discuss mutual expectations and needs and reflect upon new job profiles and qualification pathways (T2); networking PhD students who are integrating digital methods and digital transformation in historical scholarship in order to strengthen their position in the field (T3); offering short-term postdoc fellowships to share research data from their PhD or other projects, demonstrating the publication of research data (T4); and communicating the digital transformation in historical scholarship and resulting shifts in our perspectives to the larger public (T5).

TA5-M5-T1: Clio guides as communication hub: Developing the Clio Guide into an interactive communication hub for discussion on the digital transformation (M1), information on data and methods (M2), legal and ethical issues (M3) and quality and review criteria (M4) on research data and data-driven research and operation of the website.

TA5-M5-T2: Interdisciplinary summer schools: organising, with TA4, interdisciplinary summer schools for MA students, early career researchers and practitioners from academic research, memory institutions, public history and citizen science projects as a platform for shared learning, negotiating demands and discussing practices, job profiles and new career paths across institutions.

TA5-M5-T3: Digital history graduate network: Establishing a network of PhD students in digital history and their supervisors from Germany and abroad through annual retreats and other forms of exchange and collaboration (e.g. co-supervision) to promote the innovative impact of the next generation of researchers in the conceptual and methodological development of historical scholarship in the context of digitisation.

TA5-M5-T4: FAIR-data fellowships: Organisation of short-term fellowships for postdocs for the publication of high-quality research data from their dissertation or postdoc projects at institutions providing the fellows with the necessary research data management support.

TA5-M5-T5: 4Memory podcast: In cooperation with all other TAs, the conception and production of a regular podcast series on the transformation of historical studies in the context of digitisation, aimed primarily at interested members of the public.

Contributing members: *Co-applicants:* BSB, DM, h_da, HI, HU, IEG, LABW, VHD; *Participants:* BUW-IGP, Clio-online, DSM, FB-Gth, HAB, RHohls/HU, SBB-PK, SUB, UBI, UL, UP-DH, VHD-DGW, ZZF

HU organises T1 and T3, VHD T2, T4 and T5. LABW, HI, BSB, and ZZF organise/host summer schools in first funding period. UBI, HU, UP-DH, UL and BUW support the Digital History Graduate network, IEG, HI, DM and HAB administer the "FAIR-data fellowships".

5.5.3 Milestones and Deliverables

Overview table on major milestones (MS) and deliverables (D) and the scheduled completion (project month).

Type	No	Title	Month
MS	1a	Expert group in session	6
MS	1b	First international workshop on epistemological implications of digital transformation	24
D	**I**	**Hybrid book publication on the consequences of datafication in the field of history**	**48**
MS	2a	Initial dataset from survey of digital methods in current research projects	12
MS	2b	First “4Memory History Hackathon” taken place	18
D	**II**	**Clio Guides on research data and digital methods**	**36**
MS	3a	Expert group in session	8
MS	3b	First multimedia story published	24
D	**III**	**Clio Guides FAQ on legal and ethical issues**	**42**
MS	4a	Review guidelines for research data and data-driven publications	18
MS	4b	List of requirements for a research data review module	30
D	**IV**	**Research data review module at H-Soz-Kult**	**42**
MS	5a	First Interdisciplinary Summer School taken place	12
MS	5b	First FAIR-data fellowships completed	18
D	**V**	**First version of Clio Guides communication hub**	**24**

5.5.4 Cooperation with other task areas and cross-area dependencies

TA5 works closely with other TAs: to understand digital transformation in our fields, it builds on TAs 1-4 and feeds insights back into them. This applies, e.g., to the survey of the use of research data and data-driven methods, which will identify fields of action to all TAs. TA5 cooperates with TA1 to define evaluation criteria for reviewing research data. It works with TA2 on new forms of publication and authorship, where incorporating metadata is a key issue. Authority files developed there are also used in the MethodLab, where TA3’s data and services are used, especially from the 4Memory Data Lab. New methods link TA5 and TA4, since their use requires the data literacy developed and disseminated by the latter, a process to which TA5 contributes. Several events are organised jointly with TA4, providing cost-effectiveness and bringing together their perspectives. Through its community involvement, TA5 serves TA6 as an important communication hub to the historically oriented humanities, and it cooperates with TA6 in the architecture and maintenance of the corresponding technical infrastructure. TA5 also contributes to 4Memory’s integration into the international research landscape by seeking exchange with the historical community abroad (with TA6), for example by organising conferences in other countries and inviting colleagues from abroad, particularly from neighbouring countries, to participate in the planned graduate network.

5.5.5 Quality Management and Risk Analysis

TA5’s measures are primarily communicative and face various risks, especially since they aim to speak to the entire community and involve it in the resulting discourse. Choosing the wrong topics or taking the wrong approach could diminish community support; reluctance – or inability – to integrate new methods on the part of the community may also pose a risk. Insufficient visibility for 4Memory’s efforts could reduce the pool of qualified candidates for the hackathons and

fellowships we offer; alternatively, too much communication might cause information fatigue and resistance to participating in our activities. Ultimately, poor design for the Clio-Guides or the research data review module could jeopardise TA5's objectives. To counter these risks, TA5 relies on an agile and modular approach that sees community input as the driving force and keeps the measures scalable. TA5's development and community reception will be regularly reviewed, allowing us to identify and respond to problems: surveys and evaluations will be set up, access rates to the website will be recorded and user feedback for the platform will be obtained. Candidates for collaborative research projects, fellowships and participation in the research data review will be proactively identified and approached. Above all, the diversity of the participants will ensure a broad and solid anchoring in the community, helping to make potential risks visible as early as possible.

5.6 Task Area 6: Participation and Steering

5.6.1 Description of the Task Area

TA6 enables 4Memory's collaborative work – such as handling administration, building infrastructure and setting policies – and contributes to its content, acting as an overarching layer to foster synergies, facilitate knowledge exchange, encourage participatory decision-making, maintain relationships and manage engagement in the NFDI and our community. The concepts "Participation and Steering" shape the governance structure and guide community integration (see 3.4), defining leitmotifs for coordinating and collaborating with co-applicants and participants. Through this **architecture of participation and steering** (M1), TA6 brings together spokespersons and TA coordinators, providing a forum for exchange and strategic decision-making. Academic coordination (M2) guides and enriches contributions to cross-cutting topics in 4Memory, with other consortia, in the NFDI and beyond. Our innovation framework (M3) fosters

the innovative potential in our large community, providing organisational, financial and infrastructural support. Communication and community management (M4) brings together efforts to connect 4Memory to its community. Administrative services and procedures are key parts of the consortium (M5): TA6's **dual focus on administration and content** ensures these measures' integration with 4Memory's research efforts.

TA6's work programme is conducted strategically with the co-spokespersons and operationally with a **4Memory Coordination Office (4MCO)** as the core entities to establish, supervise, manage, reflect upon and improve key aspects of the consortium. The 4MCO includes:

- The **Managing Director (MD)** coordinates steering and strategic planning, monitors TA activities, supervises evaluation and quality-control, handles engagement with the NFDI, undertakes overall controlling and provides a single point of contact for the community.
- The **Communication and Community Officer (CCO)** serves the consortium – supporting outreach and representation activities – and the community, providing enriched information and inreach opportunities through transparent information and active engagement.
- The **Financial Officer (FO)** manages the budget, the transfer and allocation of resources, procurement, contract management, financial reports and is consulted on financial issues.
- The **Project and Process Officer (PPO)** develops and formalises 4Memory's procedures, including governance processes, participation rules, consortial policies, operational models, communication workflows and quality assessment. A project management expert, the PPO supports monitoring, controlling, reporting and data management.
- The **Technical Officer (TO)** consults on strategic planning, ensures inter-task-area interoperability and contributes to cross-cutting topics on technical services in the NFDI. The TO contributes to the consortial information infrastructure and web portal, supervises technical documentation and evaluates components integrated into 4Memory.

The 4MCO's composition gives it the necessary personnel and resources for its academic, technical, financial and project-management roles: more than simply handling funds and administration, the 4MCO shapes 4Memory's *content*, supervises the consortium's *development* and contributes to meeting its *objectives*. Focused on implementing and improving processes – through tasks that arise continuously – TA6's measures run throughout the project's duration.

TA6's coordinating role means it contributes to all LINKAGE objectives. It particularly fosters ***L****inking Research, Memory Institutions and Infrastructures* through direct collaboration in the consortium and the networks of the (co-)applicant institutions. Its community management efforts will extend the ***N****etwork of Historically Oriented Research Communities*. Its contribution to the NFDI ***I****ntegrates Historical Source Criticism into Data Services* and shapes the ***K****nowledge Order for the Digital Future of the Past*.

5.6.2 Measures

Measure 1: Participation and steering architecture

This measure formally integrates the community and ensures the consortium's structural integrity and operational sustainability, including its RDM infrastructure and services. The community is integrated through formal procedures developed by the 4MCO and monitored jointly with the community (T1), including proposals, partnerships, decision-making and collaborations to advance or complement the work programme. To enable participatory yet efficient steering, the spokespersons are continuously involved in strategy through formal, transparent processes documented and communicated within the consortium (T2). Policy development – including compliance management – will lead to standardised workflows and fewer case-by-case analyses, ensuring greater equality and inclusion (T3). The consortium's sustainable operational model (T4) joins different models, reflecting various services and hosting institutions. Sustainable models for services derived from research projects, such as data collections without institutional curation, are also included (a need expressed in our problem stories.)[19] Our community has key roles in ongoing needs assessment (T5).

TA6-M1-T1: Community integration and participation procedures: These procedures enable the active collaboration, strategic involvement and the on- and offboarding of initiatives, projects and institutions to and from the consortial network and the infrastructure ecosystem.

TA6-M1-T2: Steering the consortium and strategic planning: The governance bodies and their respective participation and decision-making processes are set up, continuously monitored and improved. Co-spokespersons and the MD are accountable for ongoing governance.

TA6-M1-T3: Policy development and compliance management: Participation and steering are supported by consortial policies and compliance-management procedures. Initial policies include the consortial agreement, project charter, code of conduct, mission statement, equal opportunity actions, barrier-free information access and conflict management.

TA6-M1-T4: Sustainable operational model: The operational models of the consortium's services will be evaluated and elaborated to enable upscaling from local to national/international or institutional to disciplinary levels, and a **common framework for 4Memory services** is defined and aligned with the framework developed in the NFDI.

TA6-M1-T5: Evaluation and Quality Assurance: The Advisory Board and the Community Forum will have the mandate and instruments to assess 4Memory's performance according to consortial objectives and the needs of the community, which will be surveyed regularly by the participants. Quality management in all TAs is aligned to ensure common principles.

[19] PS 33, 47, 48, 50, 57, 62, 64, 65, 67, 69, 70, 75.

Contributing members: *(Co-)applicants:* BSB, DM, FIZ, h_da, HI, HU, IEG, LABW, MLU-HistData, UT, VHD

The 4MCO at the IEG is responsible for the execution of the tasks, providing the operational layer; all (co-)spokespersons form the strategic layer, exercising their rights and duties.

Measure 2: Academic coordination

Academic coordination facilitates collaboration among the TAs and initiates interactions among the consortial members, ensuring technical, organisational and social interoperability. The scope of coordinating activities includes organising and contributing to teamwork within the consortium and the community (T1). The technical coordination ensures the interoperability and assessment of services on the technological level (T2). Ethical and legal issues[20] are addressed both specifically (T3) and in general (T4). The transfer of the consortium's key results (T5) is subject to multiple activities and publication formats, overseen by the editorial board.

TA6-M2-T1: Coordination of historical research data management: The MD, supported by the 4MCO and the Steering Committee, oversees all consortial activities to identify synergies and initiate collaborations while avoiding redundancy. These aims are fulfilled through regular updates, work-programme reviews, a common vocabulary and an open information flow.

TA6-M2-T2: Technical Coordination: The TO centrally supports interoperability among task areas (components and infrastructure) and NFDI-wide and/or generic data management services. Furthermore, the evaluation of external technical components (cross-consortial collaboration or community integration), the implementation of common procedures and standards and the technical documentation are supported and supervised by the TO.

TA6-M2-T3: Legal expertise: Legal experts from FIZ support the 4MCO to identify community needs from the use cases in all task areas. The experts provide legal guidance while ensuring close collaboration within the NFDI to complement other services and cross cutting activities.

TA6-M2-T4: Special Committee on Research Ethics: A group of international experts supervises the 4Memory consortium on ethical questions. The committee members complement perspectives and expertise in 4Memory, give advice on issues arising from the measures related to ethical questions, evaluate these activities and set impulses for future work.

TA6-M2-T5: Publication and transfer of results: The 4Memory Editorial Board ensures quality management and review processes for all 4Memory-publications of all TAs. The transfer of 4Memory's results into politics and society are initiated, guided and supported.

Contributing members: *(Co-)applicants:* BSB, DM, FIZ, h_da, HI, HU, IEG, LABW, MLU-HistData, UT, VdA, VHD; *Participants:* ZDV

[20] Legal issues recur among our problem stories: PS 9, 24, 30, 45, 48, 50, 51, 52, 55, 70, 74, 79, 81, 84.

The 4MCO based at the IEG is responsible for all tasks. The computing centre ZDV undertakes technical coordination (T2). Experts from FIZ provide legal guidance and foster collaboration within the NFDI (T3). The Special Committee on Research Ethics will be organised by IEG, which invites external experts to participate. The following areas will be addressed: algorithmic methods and artificial intelligence; data provenance; non-western perspectives; privacy and sensitive data (T4). Initially composed of members drawn from the (co-)applicant institutions, the Editorial Board will be expanded with representatives of our participants (T5).

Measure 3: Innovation framework and engagement

4Memory will incorporate and support cutting-edge RDM methods, products and processes in its community, which are constantly being developed; at the same time, the consortium must maintain consistent liaison with relevant national and international networks and perspectives. For innovations arising in the first project phase, parts of the budget are reserved in an incubator fund for which awarding procedures are established (T1). Apart from the 4MCO, many participants will serve as community agents embedded in their respective (sub-)communities (T2) – including societies and associations; citizen-science organisations; state RDM initiatives; humanities data centres; and collaborative projects and centres (CRCs, GRKs, Merian Centres) – keeping us up to date on the latest developments in our community. The scope of coordinating activities includes cross-consortial co-operations within the NFDI and contributions to and benefits from cross-cutting topics (T3). Networking on a global level takes place through liaisons, joint projects and knowledge exchange with international partners (T4).

TA6-M3-T1: Incubator for complementary and innovative actions: Strategically important project extensions will evolve through cooperation with the NFDI and with the growing 4Memory community in ways currently unforeseeable. An incubator for complementary and innovative actions ensures the flexibility and innovative potential of the consortium. The funds are centrally managed by the applicant institution; the Steering Committee decides on the allocation to TAs.

TA6-M3-T2: Participant and partner liaison: 4Memory establishes community agents in sub-communities, specific entities and interest groups: they aid the coordination of our broad-based consortium and encourage open, two-way flows of information.

TA6-M3-T3: Engagement in the NFDI and cross-cutting topics: The contributions to NFDI cross-cutting topics via projects, working groups and events from 4Memory are centrally coordinated to ensure efficient expert involvement and effective knowledge exchange. The 4MCO fosters collaborations with other NFDI consortia and supports participants' activities.

TA6-M2-T4: Internationalisation in research and infrastructure: TA6 supports all TAs' contacts with international partners and the development of the consortium's internationalisation strategy (see 3.3). A special committee including representatives of participant and partner

institutions representing several world regions advise the coordinators of the consortium's international activities and provide further outreach channels.

Contributing members: *(Co-)applicants:* BSB, DM, FIZ, h_da, HI, HU, IEG, LABW, MLU-HistData, UT, VHD; *Participants:* MWF Delhi, RHohls/HU

All spokespersons decide on disbursement of incubator funds (T1). The 4MCO (IEG) organises participant and partner liaison (T2). All co-applicants are involved in all tasks and accountable for NFDI engagement (T3). The internationalisation task force (T4) initially consists of the (Co-)applicants as well as RHohls/HU and MWF Delhi with their international networks.

Measure 4: Communication and community management

4Memory's communication and community management activities will liaison with the community through meaningful outreach and inreach. Their guiding aims are establishing two-way communication and treating every request from the community as a chance to support and advance historical research. The establishment of an internal and external single point of contact (T1) – both technically and in the person of the MD – enables clear, consistent communication. Through organisational workflows and digital tools, the 4MCO's officers ensure constant information flows, forming the core of the communication strategy (T2). Public activities focus both on inreach and outreach (T3), offering information to and input from the community. The 4Memory web portal will be a knowledge base and guidebook for the community (T4). The key consortial events – the annual community meeting and the advisory board – are centrally supported (T5) and reflected in the budget of direct costs.

TA6-M4-T1: Provide a central helpdesk: As the consortium's primary point of contact the MD is supported with a centrally managed helpdesk system. A clear contact point encourages requests and the 4MCO will avoid informational bottlenecks by implementing procedures for intermediation and forwarding requests to the task areas, working groups and experts. Thus, the helpdesk ensures the timely, reliable processing of requests.

TA6-M4-T2: Develop and implement a communications strategy: The consortium relies on many communication flows: inner- and inter-consortial, community, NFDI-wide, international and public. A communications strategy outlines relevant channels, workflows and policies.

TA6-M4-T3: Coordinate inreach and outreach activities: 4Memory's involvement with its community builds on regular participation and presence in community-organised workshops and conferences, the consortium's own events and supporting existing initiatives. Relevant calls and events are prepared. Contributions are supported by target-group-related information materials.

TA6-M4-T4: Design, implement, coordinate and maintain the 4Memory portal: A web portal serves as a central information hub, documenting the consortium and its services. Rather than a

one-stop-shop, it provides a comprehensive guide to 4Memory's modular services ecosystem. Design, development, implementation and maintenance are incorporated into the work plan.

TA6-M4-T5: Consortial event management: Central event management enables professional support for consortium-wide and governance events (invitation, communication, reporting, travel grants). (Co-)applicants receive centrally managed funds when hosting and conducting consortial, cross-area or cross-consortial events.

Contributing members: *(Co-)applicants:* BSB, DM, FIZ, h_da, HI, HU, IEG, LABW, MLU-HistData, UT, VHD; *Participants:* ZDV

The 4MCO at the IEG is responsible for all tasks, all co-applicants serve as information multipliers to the subcommunities and contribute to outreach and inreach activities. ZDV is consulted for the design, implementation and maintenance of the 4Memory web portal.

Measure 5: Administration and procedures

4Memory's administration aims for professional management and a comprehensive service portfolio to allow operational staff to efficiently implement the work programme and provide the (co-)spokespersons with effective, trustworthy and accountable workflows. The FO enables the 4MCO to set up reliable, responsive financial management (T1). The 4MCO's professional project and process management eases controlling and enables change management (T2). A cloud-based project information and communication infrastructure (T3) meets the needs for digital tools in day-to-day work while enabling a secure data and information flow. 4Memory sets high standards for its own data management and documentation (T4). Centralised monitoring of funding opportunities (T5) will initiate collaborations on joint project proposals.

TA6-M3-T1: Financial management and financial controlling: The FO manages financial flows, budget transfers, resource allocation and procurement and contracts of outsourced services. The financial reports and the incubator investments (TA6-M3-T1) are overseen.

TA6-M3-T2: Progress reporting and project controlling: The 4MCO provides professional project management processes, including progress monitoring, roadmap development, review of reports, responsibility distribution, performance evaluation and change management. The information gathered in controlling will also be incorporated in a public transparency report.

TA6-M3-T3: Infrastructure for documentation, communication and project management: The consortium and the community will be supported by a comprehensive set of tools bundled in 4Memory's information and communication cloud infrastructure. Task management, project controlling, communication channels, general administration and data management rely on established tools, including file-sharing, issue tracking, mailing lists, video conferencing, surveys, wikis, messengers, text editing, version control systems and web servers. GDPR-compliance and account management enable a smooth and secure environment.

TA6-M3-T4: Documentation and data management: 4Memory will produce substantial data, documents and materials. Centrally supported data management and documentation workflows ensure results are stored and published according to the standards of our RDM strategy.

TA6-M3-T5: Monitoring for joint proposals and funding opportunities: A curated list of funding calls and grants relevant to 4Memory's community, of needs identified and assessed by the community and of project ideas derived from 4Memory activities will facilitate collaborative efforts on future grant proposals and ease finding project partnerships.

Contributing members: *(Co-)applicants:* BSB, DM, FIZ, h_da, HI, HU, IEG, LABW, MLU-HistData, UT, VHD; *Participants:* ZDV

The 4MCO at the IEG is responsible for all tasks; all co-applicants contribute to central administration through dedicated administrative resources in the respective TAs. The ZDV's self-contributions to the information and communication infrastructure address crucial needs.

5.6.3 Deliverables and milestones

Overview table on major milestones (MS) and deliverables (D) and the scheduled completion (project month).

Type	No	Title	Month
MS	1a	All governance bodies and participation procedures established	12
MS	1b	Initial set of consortial policies adopted	12
MS	1c	Mid-term chronicle of incubator projects	30
D	**I**	**Sustainable operational model**	**54**
MS	2a	Internationalisation strategy	12
MS	2b	Guidelines and framework for technical documentation	18
D	**II**	**Established participation in NFDI and contributions to cross-cutting topics**	**24**
MS	3a	Project infrastructure setup	6
MS	3b	Public documentation	60
D	**III**	**Public transparency report**	**30**
MS	4a	Open community assessment on implemented and planned strategies and services	24
MS	4b	Community growth (project-based, national partners, international network)	48
D	**IV**	**4Memory portal launch including accompanying services**	**18**

5.6.4 Cooperation with other task areas and cross-area dependencies

Via the 4MCO, TA6 centrally manages overarching tasks that are collaboratively carried out by 4Memory's members. Coordination is enabled through a) co-applicant involvement in all strategic tasks and b) the centralised core team at the applicant institution, complemented by decentralised administrative tasks handled by co-applicants. Designed as a joint effort of all (co-)applicants, this TA serves the others, allowing them to focus on their work programmes and providing capacities for overarching, inter-consortial and NFDI-wide activities. TA6 is thus a bridge for cross-area topics and to activities beyond 4Memory. While each TA has direct connections to the community through its work programme, TA6 ensures the broad involvement of 4Memory's community through inclusive, far-reaching community management; it also cooperates with other TAs on specific challenges regarding communication infrastructure (TA1, TA5), community engagement (TA1, TA2, TA5), and operational models (TA3, TA4).

5.6.5 Quality Management and Risk Analysis

To foster collaboration and engagement of all consortium members the large, diverse 4Memory consortium relies on the established cooperation of the (co-)applicants and deep integration in many parts of our community (through joint projects, networks, working-groups, and other forms of collaboration). Policies, such as the code of conduct, will be developed to avoid conflicts while a designated ombudsperson (member of the Academic Advisory Board) will have the responsibility (and resources) for managing problems that may nevertheless arise. TA6 seeks to retain flexibility and openness to innovation through carefully calibrated onboarding processes, its incubator fund and additional joint project proposals from among its membership. The co-applicants have developed effective working relationships and significant trust during the lengthy application phase, which will be maintained through distributions of responsibility, effective processes, transparent documentation and multiple bidirectional communication channels. Potential fluctuations in personnel will be covered through committed self-contributions from the applicant institution and supported by the co-applicants.

ⓘ *Pages 117-121 contained details regarding project finances and have been removed from this public version of the funding application.*

List of Abbreviations

4MCO	4Memory Coordination Office
AdW-Göttingen	Akademie der Wissenschaften zu Göttingen
AdW-Mainz	Akademie der Wissenschaften und der Literatur Mainz
AHiG	AG Historische Grundwissenschaft / Nachwuchsnetzwerk Historische Grundwissenschaften
AIP	Leibniz-Institut für Astrophysik Potsdam
AK-PF	Arbeitskreis Provenienzforschung e.V., Berlin
AKucz/AdW	Prof. Dr. Andreas Kuczera, Akademie der Wissenschaften und der Literatur Mainz / Technische Hochschule Mittelhessen
AMIGJ	Arye Maimon-Institut für Geschichte der Juden, Universität Trier
AP-D	Archivportal-D, German Archives Portal
ArchivSCH	Archivschule Marburg – Hochschule für Archivwissenschaft
BA	Bundesarchiv
BAdW	Bayerische Akademie der Wissenschaften, München
BASE	Bielefeld Academic Search Engine
BBF	Bibliothek für Bildungsgeschichtliche Forschung, Berlin
BERD@NFDI	NFDI for Business, Economic and Related Data
BerHis	berlinHistory e.V., Berlin
BfA	Bundesagentur für Arbeit
BSB	Bayerische Staatsbibliothek München
OIB	Prof. Dr. Birgit Schäbler, Orient Institut Beirut
BUW	Bergische Universität Wuppertal
BUW-DH	Bergische Universität Wuppertal: Digital Humanities
BUW-IGP	Bergische Universität Wuppertal: Institut für Grundlagenforschung zur Philosophiegeschichte
BUW-IZED	Bergische Universität Wuppertal: Interdisziplinäres Zentrum für Editions- und Dokumentwissenschaft
C2DH	Luxembourg Centre for Contemporary and Digital History
CARE	Collective benefit, authority to control, responsibilty, ethics
CdV	Coding da Vinci
CERES	Ruhr-Universität Bochum: Centrum für Religionswissenschaftliche Studien
CERL	Consortium of European Research Libraries
CIDOC	Comité international pour la documentation (ICOM)
Clio-online	Clio-online - Historisches Fachinformationssystem e.V.
CompGen	Verein für Computergenealogie e.V., Köln
DARIAH-DE	Digital Research Infrastructure for the Arts and Humanities (German)
DARIAH-EU	Digital Research Infrastructure for the Arts and Humanities (ERIC)
DBV	Deutscher Bibliotheksverband
DDK	Deutsches Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg
DGO	Deutsche Gesellschaft für Osteuropakunde e. V.
DGPhil	Deutsche Gesellschaft für Philosophie e.V.
DGV	Deutsche Gesellschaft für Völkerkunde e.V.
DHd	Verband Digital Humanities im deutschsprachigen Raum e.V.

DHI London	Prof. Dr. Christina von Hodenberg, Deutsches Historisches Institut London
DHI Moskau	Dr. Sandra Dahlke, Deutsches Historisches Institut Moskau
DHI Paris	Prof. Dr. Thomas Maissen, Deutsches Historisches Institut Paris
DHI Rom	Prof. Dr. Martin Baumeister, Deutsches Historisches Institut Rom
DHI Washington	Prof. Dr. Simone Lässig, Deutsches Historisches Institut Washington
digiCULT	digiCULT-Verbund eG, Kiel
DIJ Tokyo	Deutsches Institut für Japanstudien Tokio
DINI	Deutsche Initiative für Netzwerkinformation e. V.
DIPF	Leibniz-Institut für Bildungsforschung und Bildungsinformation
DL	Data Literacy
DM	Deutsches Museum
DME	Data Modeling Environment
DMP	Data Management Plan
DNB	Deutsche Nationalbibliothek
DSM	Deutsches Schifffahrtsmuseum – Leibniz-Institut für Maritime Geschichte, Bremerhaven
DVRW	Deutsche Vereinigung für Religionswissenschaft e. V.
EHRI	European Holocaust Research Infrastructure
EMIW	Emil-Frank-Institut, Wittlich
EOSC	European Open Science Cloud
ERIC	European Research Infrastructure Consortium
ExClu AM	Exzellenzcluster “Africa Multiple”, Universität Bayreuth
ExClu UWA	Cluster Understanding Written Artefacts, Universität Hamburg
FAIRmat	FAIR Data Infrastructure for Condensed-Matter Physics and the Chemical Physics of Solids
FAU	Friedrich-Alexander-Universität Erlangen-Nürnberg
FB Gth	Forschungsbibliothek Gotha
fdm.nrw	Landesinitiative für Forschungsdatenmanagement – NRW
FID	Fachinformationsdienst
FIZ	Leibniz-Institut für Informationsinfrastruktur
FSU	Friedrich-Schiller-Universität Jena
FSU-FDM	Friedrich-Schiller-Universität Jena: Kontaktstelle Forschungsdatenmanagement
FSU-HI	Friedrich-Schiller-Universität Jena: Historisches Institut
FSU-II	Friedrich-Schiller-Universität Jena: Institut für Informatik
FSU-MEPHisto	Friedrich-Schiller-Universität Jena: Arbeitsgruppe MEPHisto – Digitale Modelle, Prozesse und Erklärungen in den Historischen Wissenschaften
FSU-ThULB	Friedrich-Schiller-Universität Jena: Thüringer Universitäts und Landesbibliothek
FU	Freie Universität Berlin
FU-CeDiS	Freie Universität Berlin: Center für Digitale Systeme
FU-Didaktik	Freie Universität Berlin: Arbeitsbereich „Didaktik der Geschichte“
FU-IRTG	Freie Universität Berlin: IRTG Temporalities of Future
FU-LAI	Freie Universität Berlin: Lateinamerika-Institut
FU-UB	Freie Universität Berlin: Universitätsbibliothek
FUH-Archiv	Fern Universität Hagen: Archiv des Instituts für Geschichte und Biographie

DIJ Tokyo	Prof. Dr. Franz Waldenberger, Deutsches Institut für Japanstudien, Tokyo
FZH	Forschungsstelle für Zeitgeschichte in Hamburg: Werkstatt der Erinnerung
GBV	Gemeinsamer Bibliotheksverbund
GDA	Generaldirektion der Staatlichen Archive Bayerns
GdGA	Gesamtverein der deutschen Geschichts- und Altertumsvereine e.V.
GEI	Georg-Eckert-Institut – Leibniz-Institut für Internationale Schulbuchforschung, Braunschweig
GFZ	Deutsches GeoForschungsZentrum
GNM	Germanisches Nationalmuseum – Leibniz-Forschungsmuseum für Kulturgeschichte, Nürnberg
GoFAIR	GO FAIR is an initiative that aims to implement the FAIR data principles
GSWG	Gesellschaft für Sozial- und Wirtschaftsgeschichte: Lehrstuhl für Wirtschafts- und Sozialgeschichte, Universität Regensburg
GU-UB	Goethe-Universität Frankfurt: Universitätsbibliothek
GWZO	Leibniz-Institut für Geschichte und Kultur des östlichen Europa
h_da	Hochschule für angewandte Wissenschaften Darmstadt
H-Soz-Kult	H-Soz-Kult: Kommunikation und Fachinformation für die Geschichtswissenschaften (Clio-online e.V.)
HAB	Herzog August Bibliothek Wolfenbüttel
HDU	Ruprecht-Karls-Universität Heidelberg
HDU-CATS	Ruprecht-Karls-Universität Heidelberg: Centre for Asian and Transcultural Studies
HDU-HCTS	Ruprecht-Karls-Universität Heidelberg: Heidelberg Center for Transcultural Studies
HDU-HRA	Ruprecht-Karls-Universität Heidelberg: Heidelberg Research Architecture
HDU-HSE	Ruprecht-Karls-Universität Heidelberg: Heidelberg School of Education
HDU-RC	Ruprecht-Karls-Universität Heidelberg: Research Council – Kulturelle Dynamiken in globalisierten Welten
HDU-UB	Ruprecht-Karls-Universität Heidelberg: Universitätsbibliothek
HeFDI	Hessische Forschungsdateninfrastrukturen
HI	Herder-Institut für historische Ostmitteleuropaforschung - Institut der Leibniz-Gemeinschaft, Marburg
HIS	Hamburger Institut für Sozialforschung: Soziopolis
HLA	Hessisches Landesarchiv
HU	Humboldt Universität zu Berlin
HU-CMS	Humboldt Universität zu Berlin: Computer- und Medienservice
HU-UB	Humboldt Universität zu Berlin: Universitätsbibliothek
Huygens ING	Huygens Institute for the History of the Netherlands
HVIW	Hochschulverband Informationswissenschaft
IAE	International Association of Egyptologists
IAI	Ibero-Amerikanisches Institut Berlin
ICARUS	International Centre for Archival Research
ICOM	International Council of Museums
IDE	Institut für Dokumentologie und Editorik e.V.
IEG	Leibniz-Institut für Europäische Geschichte, Mainz
IFLA	International Federation of Library Associations and Institutions
IFZ	Institut für Zeitgeschichte München-Berlin

IGDJ	Institut für die Geschichte der deutschen Juden, Hamburg
IISG	International Institute of Social History
IRS	Leibniz-Institut für Raumbezogene Sozialforschung Erkner
ISGV	Institut für Sächsische Geschichte und Volkskunde, Dresden
IStG	Institut für vergleichende Städtegeschichte Münster
KBL	Kommission für bayerische Landesgeschichte
KGD	Konferenz für Geschichtsdidaktik e.V.
KIBA	Konferenz der informations- und bibliothekswissenschaftlichen Ausbildungs- und Studiengänge
KonsortSWD	Konsortium für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften
LABW	Landesarchiv Baden-Württemberg
LIBER	Ligue des Bibliothèques Européennes de Recherche
LMU	Ludwig-Maximilian-Universität München
LMU-ITGW	Ludwig-Maximilian-Universität München: IT-Gruppe Geisteswissenschaften
LoC	Library of Congress
LOD	Linked Open Data
LRZ	Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften
MADOC	Publikationsserver der UBM
mainzed	Mainzer Zentrum für Digitalität in den Geistes- und Kulturwissenschaften
MARC	Machine-Readable Cataloging
MaRDI	Mathematical Research Data Initiative
MG	Mommsen-Gesellschaft e.V., Freiburg
MGH	Monumenta Germaniae Historica, München
MLU	Martin-Luther-Universität Halle-Wittenberg
MLU-HistData	Martin-Luther-Universität Halle-Wittenberg: Historisches Datenzentrum Sachsen Anhalt
MLU-IZEA	Martin-Luther-Universität Halle-Wittenberg: Interdisziplinäres Zentrum für die Erforschung der Europäischen Aufklärung
MMZ	Moses Mendelssohn Zentrum, Potsdam
MPIWG	Max-Planck-Institut für Wissenschaftsgeschichte
MV	Mediävistenverband e.V., Frankfurt a. M.
MWF Delhi	Dr. Sebastian Schwecke, Max Weber Forum für Südasienstudien Delhi
nestor	Network of Expertise in long-term Storage and availability of digital Resources in Germany
NFDI-MatWerk	National Research Data Infrastructure for Materials Science and Engineering
NFDI4Agri	NFDI for Agricultural Sciences
NFDI4Chem	Chemistry Consortium in the NFDI
NFDI4Culture	Consortium for research data on material and immaterial cultural heritage
NFDI4DataScience	NFDI for Data Science and Artificial Intelligence
NFDI4Earth	NFDI Consortium Earth System Sciences
NFDI4Microbiota	National Research Data Infrastructure for Microbiota Research
NFDI4Objects	Research Data Infrastructure for the Material Remains of Human History
NFDI4SD	NFDI4SDSmall Disciplines
NFDIxCS	National Research Data Infrastructure for and with Computer Science
OAH	Organization of American Historians

OFU	Otto-Friedrich-Universität Bamberg: Lehrstuhl für Medieninformatik
OIB	Prof. Dr. Birgit Schäbler, Orient Institut Beirut
OPERAS	European Research Infrastructure for the development of open scholarly communication in the social sciences and humanities
OWL	Web Ontology Language
PUNCH4NFDI	Particles, Universe, NuClei and Hadrons for the NFDI
PS	4Memory Problem Stories
RDA	Research Data Alliance
RDF	Resource Description Framework
RDM	Research Data Management
RDMO	Research Data Management Organiser
RESILIENCE	REligious Studies Infrastructure: tooLs, Expert, conNections and CEnters in Europe
RHohls/HU	Prof. Dr. Rüdiger Hohls, Institut für Geschichtswissenschaften, HU Berlin
RRCHNM	Roy Rosenzweig Center for History and New Media
SAW	Sächsische Akademie der Wissenschaften zu Leipzig
SBB-PK	Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
MWF Delhi	Dr. Sebastian Schwecke, Max Weber Forum für Südasienstudien, Delhi
SPARQL	SPARQL Protocol And RDF Query Language
SSHOC	Social Sciences and Humanities Open Cloud
SIS	Specialised Information Services
SLUB	Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden
SoQuZ	AG Sozialdaten als Quellen der Zeitgeschichte, Trier, London, Mannheim, Halle
SPK	Stiftung Preußischer Kulturbesitz
SUB	Niedersächsische Staats- und Universitätsbibliothek Göttingen
Text+	Text+: Language-and Text-Based Research Data Infrastructure
THK	Technische Hochschule Köln: Institute of Information Science
TIB	Technische Informationsbibliothek (Hannover)
UA	Universität Augsburg
UBI	Universität Bielefeld: Fakultät für Geschichtswissenschaft, Philosophie und Theologie
UBT	Universität Bayreuth
UE	Universität Erfurt: Professur für Globalgeschichte des 19. Jahrhunderts
UE-FZG	Universität Erfurt: Forschungszentrum Gotha
UHH	Universität Hamburg
UHH-FDM	Universität Hamburg: Zentrum für nachhaltiges Forschungsdatenmanagement
UI	Universität Innsbruck
UK	Universität Konstanz
UL	Universität Leipzig
UL-AG	Universität Leipzig: Lehrstuhl für Alte Geschichte, Historisches Seminar;
UL-DHLab	Universität Leipzig: Digital Humanities Lab
UL-GESI	Universität Leipzig: Global and European Studies Institute
UL-GSGAS	Universität Leipzig: Graduate School Global and Area Studies
UL-RCG	Universität Leipzig: Leipzig Research Center Global Dynamics
UL-UB	Universität Leipzig: Universitätsbibliothek Leipzig

UBI	Universität Bielefeld
UBM	Universitäts- und Landesbibliothek Mannheim
UoC	University of Cambridge
UP-DH	Universität Passau: Lehrstuhl für Digital Humanities
UPVM3	Université Paul Valéry Montpellier 3
UR	Universität Rostock: Institut für Medienforschung
UT	Universität Trier
UT-SDCH	Universität Trier, Working Group Social Data and Contemporary History
UT-SEAL	Universität Trier: Forschungs- und Dokumentationsstelle SEAL: Strukturen und Erinnerung. Angewandte Geschichtswissenschaft und digitale Lehre
UT-SeS	Universität Trier: Servicezentrum eSciences
UT-TRANSMARE	Trierer Institut zur Erforschung des Transfers von Menschen, Gütern und Ideen von der Antike bis zur Gegenwart
UzK	Universität zu Köln
UzK-CCeH	Universität zu Köln: Cologne Center for eHumanities
UzK-DCH	Universität zu Köln: Data Center for the Humanities
UzK-FNZ	Universität zu Köln: Lehrstuhl für die Geschichte der Frühen Neuzeit, Historisches Institut
UzK-ZfL	Universität zu Köln: Zentrum für LehrerInnenbildung
VAR	Vereinigte Adelsarchive im Rheinland e.V.
VdA	Verband deutscher Archivarinnen und Archivare e.V.
VDK	Verband Deutscher Kunsthistoriker e.V.
VGD	Verband der Geschichtslehrer Deutschlands e. V.
VHD	Verband der Historiker und Historikerinnen Deutschlands
VHD-DGW	Verband der Historiker und Historikerinnen Deutschlands / AG Digitale Geschichtswissenschaft
VIAF	Virtual International Authority File
VOH	Verband der Osteuropahistorikerinnen und -historiker e. V.
VZG	Verbundzentrale des GBV, Göttingen
WA-VfS	Wirtschaftshistorischer Ausschuss, Verein für Socialpolitik
WIAG	Wissens-Aggregator Mittelalter und Frühe Neuzeit
ZBW	Leibniz-Informationszentrum Wirtschaft, Kiel
ZDV	Zentrum für Datenverarbeitung (ZDV), Johannes Gutenberg-Universität Mainz
ZDV-T	Zentrum für Datenverarbeitung, Universität Tübingen
ZZF	Leibniz-Zentrum für Zeithistorische Forschung Potsdam

Appendix

1 Bibliography and list of references

1.1. Bibliography

→ *PS: The 4Memory Problem Stories are available here:* https://4memory.de/problem-stories-overview

→ *Sources written by members of the consortium are highlighted in bold.*

AG FID. 2018. “Gemeinsame Stellungnahme der FID leitenden Einrichtungen und der AG Fachinformationsdienste (AG FID) zur Einbindung von Fachinformationsdiensten in den Aufbau einer nationalen Forschungsdateninfrastruktur.” https://wikis.sub.uni-hamburg.de/webis/images/a/a1/AG_FID_zu_NFDI.pdf.

AHA, American Historical Association. 2015. “Guidelines for the Professional Evaluation of Digital Scholarship by Historians.” https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians.

Andorfer, Peter. 2015. “Forschungsdaten in den (Digitalen) Geisteswissenschaften. Versuch einer Konkretisierung,” DARIAH-DE Working Papers, 14. urn:nbn:de:gbv:7-dariah-2015-7-2.

Arguing with Digital History working group. 2017. “Digital History & Argument.” White Paper. Roy Rosenzweig Center for History and New Media. https://rrchnm.org/argument-white-paper/.

Auer, Sören, Viktor Kovtun, Manuel Prinz, Anna Kasprzik, Markus Stocker, and Maria Esther Vidal. 2018. “Towards a Knowledge Graph for Science.” In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics - WIMS ’18*, 1–6. Novi Sad, Serbia: ACM Press. https://doi.org/10.1145/3227609.3227689.

Balzer, Detlev, **Barbara K. Fischer, Jürgen Kett**, Susanne Laux, Jens M. Lill, Jutta Lindenthal, Mathias Manecke, **Martha Rosenkötter**, and **Axel Vitzthum**. 2019. “Das Projekt ‘GND für Kulturdaten’ (GND4C).” *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 6 (4): 59–97. https://doi.org/10.5282/o-bib/2019H4S59-97.

Barats, Christine, Valérie Schafer, and Andreas Fickers. 2020. “Fading Away... The Challenge of Sustainability in Digital Studies,” Digital Humanities Quarterly, 14 (3). http://www.digitalhumanities.org/dhq/vol/14/3/000484/000484.html.

Blaney, Jonathan. 2021. *Doing Digital History: A Beginner’s Guide to Working with Text As Data*. Manchester University Press.

Borgman, Christine L. 2007. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, Mass: MIT Press.

Borgman, Christine L. 2012. “The Conundrum of Sharing Research Data.” *Journal of the American Society for Information Science and Technology* 63 (6): 1059–78. https://doi.org/10.1002/asi.22634.

BSB. 2018. “Forschungsdateninfrastruktur für die Geschichtswissenschaft. Positionspapier der Bayerischen Staatsbibliothek.” https://link.bsb-muenchen.de/385.

Brünger-Weilandt, Sabine, Kai-Christian Bruhn, Alexandra W. Busch, Erhard Hinrichs, Gerald Maier, **Johannes Paulmann**, Andrea Rapp, et al. 2020. 'Memorandum of Understanding by NFDI Initiatives from the Humanities and Cultural Studies', September. https://doi.org/10.5281/zenodo.4045000.

Buddenbohm, Stefan, Nathanael Cretin, Elly Dijk, Bertrand Gaiffe, Maaike de Jong, Nathalie Le Tellier-Becquart, and Jean-Luc Minel. 2016. "State of the Art Report on Open Access Publishing of Research Data in the Humanities." Report. DARIAH. https://halshs.archives-ouvertes.fr/halshs-01357208.

bwFDM. 2015. "BwFDM-Communities - Datenseite Aller User Storys." BwFDM-Communities. 2015. https://bwfdm.scc.kit.edu/cgi-bin/daten/.

Cremer, Fabian, Lisa Klaffki, and Timo Steyer. 2018. "Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften." *o-bib. Das offene Bibliotheksjournal / herausgegeben vom VDB* 5 (2): 142–62. https://doi.org/10.5282/o-bib/2018H2S142-162.

Cremer, Fabian, Lisa Klaffki, and Timo Steyer. 2019. "Redaktionssache Forschungsdaten." *Bibliothek Forschung Und Praxis* 43 (1): 118–125. https://doi.org/10.1515/bfp-2019-2018.

Cremer, Fabian, Silvia Daniel, Marina Lemaire, Katrin Moeller, Matthias Razum, and Arnošt Štanzel. 2021. "Data Meets History: A Research Data Management Strategy for the Historically Oriented Humanities." In *Cultural Sovereignty beyond the Modern State*, edited by Gregor Feindt, Bernhard Gissibl, and Johannes Paulmann, 155–78. De Gruyter. https://doi.org/10.1515/9783110679151-009.

Crymble, Adam. 2021. *Technology and the Historian: Transformations in the Digital Age*. University of Illinois Press.

Deutsche Forschungsgemeinschaft. 2019. "Guidelines for Safeguarding Good Research Practice. Code of Conduct," September. https://doi.org/10.5281/ZENODO.3923601.

Drucker, Johanna. 2009. *SpecLab: Digital Aesthetics and Projects in Speculative Computing*. Chicago: University of Chicago Press.

Drucker, Johanna. 2011. "Humanities Approaches to Graphical Display." *Digital Humanities Quarterly* 005 (1).

Edmond, Jennifer. 2016. "Will Historians Ever Have Big Data?" In *Computational History and Data-Driven Humanities*, edited by Bojan Bozic, Gavin Mendel-Gleason, Christophe Debruyne, and Declan O'Sullivan, 482:91–105. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46224-0_9.

Estermann, Beat, Adrian Gschwend, Stephan Haller, and Edward Marcos Parrales Machuca. 2020. "Basisregister und kontrollierte Vokabulare als Wegbereiter für Linked Open Data in der Schweiz." Info:eu-repo/semantics/report. Berner Fachhochschule, Institut Public Sector Transformation. https://arbor.bfh.ch/10249/.

Fickers, Andreas. 2012. 'Towards a New Digital Historicism? Doing History in the Age of Abundance'. *VIEW Journal of European Television History and Culture* 1 (1): 19–26. https://doi.org/10.18146/2213-0969.2012.jethc004.

Fickers, Andreas. 2020. "Update für die Hermeneutik. Geschichtswissenschaft auf dem Weg zur digitalen Forensik?" *Zeithistorische Forschungen* 17.1: 157–68. https://doi.org/10.14765/ZZF.DOK-1765.

Fickers, Andreas, and Tim van der Heijden. 2020. “Inside the Trading Zone: Thinkering in a Digital History Lab.” *Digital Humanities Quarterly* 014 (3). http://www.digitalhumanities.org/dhq/vol/14/3/000472/000472.html.

Föhr, Pascal. 2017. “Historische Quellenkritik im Digitalen Zeitalter.” Thesis, University of Basel. https://doi.org/10.5451/unibas-006805169.

GIDA, The Global Indigenous Data Alliance. 2019. “CARE Principles of Indigenous Data Governance.” https://www.gida-global.org/care.

Göttsche, Dirk, ed. 2019. *Memory and Postcolonial Studies: Synergies and New Directions*. Cultural Memories, vol. 9. Oxford Bern Berlin Wien: Peter Lang.

Harbeck, Matthias, and **Matthias Kaun**. 2019. “Forschungsdaten und Fachinformationsdienste – eine Bestandsaufnahme.” *Bibliothek Forschung und Praxis* 43 (1): 35–41. https://doi.org/10.1515/bfp-2019-2015.

Herr, Wiebke, Andreas C. Hofmann, and Katrin Getschmann. 2018. “Deutsche Historische Bibliografie. Präsentation, Partner Und Perspektiven. Bericht Über Die Tagung Am 25. Und 26. Oktober 2017 in München.” *Zeitschrift Für Bibliothekswesen Und Bibliographie* 65 (2–3): 130–34. https://doi.org/10.3196/18642950186523168.

Hiltmann, Torsten, Jan Keupp, **Melanie Althage**, and **Philipp Schneider**. 2021. “Digital Methods in Practice.” *Geschichte Und Gesellschaft* 47 (1): 122–56. https://doi.org/10.13109/gege.2021.47.1.122.

Hohls, Rüdiger, Thomas Meyer, Wilfried Enderle, Claudia Prinz, Torsten Schaßan, Frank M. Bischoff, Gantert Klaus, et al. 2018. “Clio-Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften,” June. https://doi.org/10.18452/19244.

Hohls, Rüdiger, Claudia Prinz, and Eva Schlotheuber, eds. 2016. *Historische Grundwissenschaften und die digitale Herausforderung*. Vol. 18. Historisches Forum. Berlin. http://edoc.hu-berlin.de/e_histfor/.

Kellendonk, Stefan, Marina Lemaire, and Lea Gerhards. 2020. “Workshop zur Implementierung von FDM-Services nach dem DIAMANT-Modell: Ein Workshopkonzept zur Anwendung des FDM-Referenzmodells DIAMANT an Forschungseinrichtungen.” 07. Universität Trier ESciences Working Papers. Trier. 10.25353/UBTR-XXXX-563D-6E00.

Kemman, Max. 2021. *Trading Zones of Digital History*. De Gruyter Oldenbourg. https://doi.org/10.1515/9783110682106.

Klump, Jens. 2017. “Data as Social Capital and the Gift Culture in Research.” *Data Science Journal* 16 (0). https://doi.org/10.5334/dsj-2017-014.

Koolen, Marijn, Jasmijn van Gorp, and Jacco van Ossenbruggen. 2019. ‘Toward a Model for Digital Tool Criticism: Reflection as Integrative Practice’. *Digital Scholarship in the Humanities* 34 (2): 368–85. https://doi.org/10.1093/llc/fqy048.

Kraus, Eva, and Matti Stöhr. 2015. “On the Way to a ‘German Historical Bibliography’: Current State and Perspectives – The DFG Project ‘Continued Cooperative Development of Historical Subject Bibliographies.’” In *Historical Bibliography as an Essential Source for Historiography*, edited by Kristina Rexová, Bernadette Cunningham, Václava Horčáková, and Věra Hanelová, 161–80. Newcastle upon Tyre: Cambridge Scholars Publishing.

Kuczera, Andreas, Thorsten Wübbena, and Thomas Kollatz. 2019. *Die Modellierung des Zweifels: Schlüsselideen und -konzepte zur graphbasierten Modellierung von Unsicherheiten ;* Vol. 4. Sonderbände der ZfdG. Wolfenbüttel: Forschungsverbund Marbach Weimar Wolfenbüttel. https://doi.org/10.17175/sb004.

Lin, Jimmy, Ian Milligan, Douglas W. Oard, Nick Ruest, and Katie Shilton. 2020. “We Could, but Should We? Ethical Considerations for Providing Access to GeoCities and Other Historical Digital Collections.” In *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, 135–44. CHIIR ’20. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3343413.3377980.

Lehmkuhl, Ursula. 2020.“Jenseits der Herausforderungen der ad-hoc-Digitalisierung; Überlegungen zur Auswirkung des digital turn auf die geschichtswissenschaftliche Lehre.” *Die Geschichtswissenschaft Im Digitalen Zeitalter*, VHD Journal, 9: 16–19.

Lemaire, Marina. 2018. “Vereinbarkeit von Forschungsprozess und Datenmanagement in den Geisteswissenschaften.” *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, December, 237-247 Seiten. https://doi.org/10.5282/O-BIB/2018H4S237-247.

Lemaire, Marina, Lea Gerhards, Stefan Kellendonk, Katarina Blask, and André Förster. 2020. “Das DIAMANT-Modell 2.0. Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-Servicelandschaft.” 05. Universität Trier ESciences Working Papers. Trier. https://doi.org/10.25353/ubtr-xxxx-f5d2-fffb.

Ma, Rongqian, and Fanghui Xiao. 2020. “Data Practices in Digital History.” *International Journal of Digital Curation* 15 (1): 21. https://doi.org/10.2218/ijdc.v15i1.597.

Maier, Gerald, Andreas Fähle, and Andreas Neuburger. 2020. “Bereitstellung, Aufbereitung, Langzeitsicherung. Funktionen der Archive in der Forschungsdateninfrastruktur,” Archivar. Zeitschrift für Archivwesen, 73 (1): 13–18.

Maik Bierwirth, Frank Oliver Glöckner, Christian Grimm, Sonja Schimmler, **Franziska Boehm**, Christian Busse, **Andreas Degkwitz**, Oliver Koepler, and Heike Neuroth. 2020. “Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung,” June. https://doi.org/10.5281/zenodo.3895209.

Milligan, Ian. 2019. *History in the Age of Abundance?: How the Web Is Transforming Historical Research*. McGill-Queen's University Press.

Minn, Gisela, and Marina Lemaire. 2017. “Forschungsdatenmanagement in den Geisteswissenschaften. Eine Planungshilfe für die Erarbeitung eines digitalen Forschungskonzepts und die Erstellung eines Datenmanagementplans.” Working Paper 03. Universität Trier ESciences Working Papers. Trier: Universität Trier. http://nbn-resolving.de/urn:nbn:de:hbz:385-10715.

Moeller, Katrin. 2019. “Standards für die Geschichtswissenschaft! Zu differenzierten Funktionen von Normdaten, Standards und Klassifikationen für die Geisteswissenschaften am Beispiel von Berufsklassifikationen.” In *Aufklärungsforschung Digital. Konzepte, Methoden, Perspektiven*, edited by Anne Purschwitz, 17–43. Halle.

Nauta, Gerhard Jan, Wietske van den Heuvel, and Stephanie Teunisse. 2017. “Report on ENUMERATE Core Survey 4.” Deliverable D4.4. Europeana DSI 2– Access to Digital Resources of European Heritage. https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI-2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf.

Oltersdorf, Jenny, and **Stefan Schmunk**. 2016. “Von Forschungsdaten und wissenschaftlichen Sammlungen.” *Bibliothek Forschung Und Praxis* 40 (2): 179–185. https://doi.org/10.1515/bfp-2016-0036.

Paulmann, Johannes, and John Carter Wood. 2020. “Die digitale Integration von Historischer Forschung, Gedächtniseinrichtungen und Infrastrukturen. Zur Konsortialinitiative 4Memory in Der NFDI,” Archivnachrichten aus Hessen, Sonderheft 2020: 77–81.

Paulmann, Johannes, John Carter Wood, and Fabian Cremer. 2020. “Linkage - Digitale Gegenwart und Zukunft historischer Forschung. Die Ziele der Konsortialinitiative 4Memory.” *Die Geschichtswissenschaft Im Digitalen Zeitalter*, VHD Journal, 9: 26–34.

Puhl, Johanna, Peter Andorfer, Mareike Höckendorff, **Stefan Schmunk**, Juliane Stiller, and Klaus Thoden. 2015. “Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften,” DARIAH-DE Working Papers, 11. http://nbn-resolving.de/%20urn:nbn:de:gbv:7-dariah-2015-4-4.

Queckbörner, Boris. 2019. “Forschungsdaten und Forschungsdatenmanagement in der Geschichtswissenschaft,” September. https://doi.org/10.18452/20460.

Rehbein, Malte. 2018. “„L‘historien de demain sera programmeur ou il ne sera pas.“ (Digitale) Geschichtswissenschaften heute und morgen.” *Digital Classics Online*, June, 23–43. https://doi.org/10.11588/dco.2017.0.48491.

RfII. 2016. “Enhancing Research Data Management: Performance through Diversity. Recommendations regarding structures, processes, and financing for research data management in Germany.” RfII – German Council for Scientific Information Infrastructures: Göttingen. http://www.rfii.de/?wpdmdl=2075.

RfII. 2020a. “Building Sustainable Data Services. RfII Discussion Paper on the Enhancement of Research Data Infrastructures.” RfII – German Council for Scientific Information Infrastructures: Göttingen. http://www.rfii.de/?wpdmdl=4422.

RfII. 2020b. “Digital competencies – urgently needed! – Recommendations on career and training prospects for the scientific labour market.” 6. RfII – German Council for Scientific Information Infrastructures: Göttingen. http://www.rfii.de/?wpdmdl=4015.

RfII. 2020c. “The Data Quality Challenge. Recommendations for Sustainable Research in the Digital Turn.” 7. RfII – German Council for Scientific Information Infrastructures: Göttingen. http://www.rfii.de/?wpdmdl=4203.

Ridsdale, Chantel, James Rothwell, Michael Smit, Hossam Ali-Hassan, Michael Bliemel, Dean Irvine, Daniel Kelley, Stan Matwin, and Bradley Wuetherick. 2015. “Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report.” Report. https://doi.org/10.13140/RG.2.1.1922.5044.

Romein, C. Annemieke, Max Kemman, Julie M. Birkholz, James Baker, Michel De Gruijter, Albert Meroño-Peñuela, Thorsten Ries, Ruben Ros, and Stefania Scagliola. 2020. “State of the Field: Digital History.” *History* 105 (365): 291–312. https://doi.org/10.1111/1468-229X.12969.

Romele, Alberto, Marta Severo, and Paolo Furia. 2020. “Digital Hermeneutics: From Interpreting with Machines to Interpretational Machines.” *AI & SOCIETY* 35 (1): 73–86. https://doi.org/10.1007/s00146-018-0856-2.

Rosenzweig, Roy. 2003. ‘Scarcity or Abundance? Preserving the Past in a Digital Era’. *The American Historical Review* 108 (3): 735–62. https://doi.org/10.1086/ahr/108.3.735.

Schlotheuber, Eva, and Johannes Paulmann. 2020. “Digitale Wissensordnung und Datenqualität: Herausforderungen, Anforderungen und Beitrag historisch arbeitender Wissenschaften,” Archivar. Zeitschrift für Archivwesen, 73 (1): 9–12.

Schöch, Christof. 2013. “Big? Smart? Clean? Messy? Data in the Humanities.” *Journal of Digital Humanities* 2 (3): 2–13.

Schroeder, Ralph, and Niels Brügger. 2017. *The Web as History : Using Web Archives to Understand the Past and the Present*. UCL Press. https://library.oapen.org/handle/20.500.12657/45665.

Schüller, Katharina, Paulina Busch, and Carina Hindinger. 2019. “Future Skills: Ein Framework für Data Literacy. Kompetenzrahmen und Forschungsbericht.” Working Paper 47. Hochschulforum Digitalisierung. Arbeitspapier. Berlin. https://doi.org/10.5281/zenodo.3349865.

Singhal, Amit. 2012. “Introducing the Knowledge Graph: Things, Not Strings.” *Google Official Blog* (blog). May 16, 2012. https://blog.google/products/search/introducing-knowledge-graph-things-not/.

Smale, Nicholas, Kathryn Unsworth, Gareth Denyer, and Daniel Barr. 2018. “The History, Advocacy and Efficacy of Data Management Plans.” *BioRxiv*, October, 443499. https://doi.org/10.1101/443499.

Staab, Steffen, and Rudi Studer. 2009. *Handbook on Ontologies*. 2nd ed. Springer Publishing.

Stalder, Felix. 2018. *The Digital Condition*. New York: John Wiley & Sons.

Tóth-Czifra, Erzsébet. 2019. “The Risk of Losing Thick Description: Data Management Challenges Arts and Humanities Face in the Evolving FAIR Data Ecosystem.” In *Digital Technology and the Practices of Humanities Research*. Open Book Publishers. https://doi.org/10.11647/OBP.0192.

VHD. 2017. “Positionspapier des Verbandes der Historiker und Historikerinnen Deutschlands (VHD) zur Schaffung nationaler Forschungsdateninfrastrukturen (NFDI).” https://www.historikerverband.de/mitteilungen/mitteilungs-details/article/positionspapier-zur-schaffung-nationaler-forschungsdateninfrastrukturen-nfdi.html.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3 (1): 160018. https://doi.org/10.1038/sdata.2016.18.

Zaagsma, Gerben. 2013. “On Digital History.” *BMGN - Low Countries Historical Review* 128.4: 3–29.

1.2. Digital Resources

→ Resources provided by members of the consortium are highlighted in bold and with asterisk.

- **Archivportal-D*** (FIZ, LABW), https://www.archivportal-d.de/
- **Arcinsys (Archivinformationssystem Hessen)*** (HLA), https://arcinsys.hessen.de/
- **ARTOS - Artikel und Rezensionen zu Ost-, Ostmittel- und Südosteuropa*** (BSB), https://www.osmikon.de/servicemenue/ueber-uns/ueber-artos/
- BASE Bielefeld Academic Research Engine (UBI), https://www.base-search.net/
- **Bavarian State Library Catalogues*** (BSB), https://www.bsb-muenchen.de/recherche-und-service/suchen-und-finden/
- **bavarikon*** (BSB and other Bavarian institutions), https://www.bavarikon.de/
- **Bayerische Landesbibliothek Online*** (BSB and other Bavarian institutions), https://www.bayerische-landesbibliothek-online.de/
- **Bibliographisch-biographische Datenbank zur Geschichte jüdischer Familien in der Region Eifel-Mosel-Hunsrück*** (EMIW, UT), http://urts81.uni-trier.de/emilfrank/bibliographiesuche2.php
- **Bildähnlichkeitssuche*** (BSB), https://bildsuche.digitale-sammlungen.de/
- **Bildarchiv der BSB (inkl. STERN-Fotoarchiv)*** (BSB), https://www.bsb-muenchen.de/sammlungen/bilder/
- **Bildkatalog des Herder-Instituts*** (HI), https://www.herder-institut.de/bildkatalog/
- **BOLSA: Die Bonner Längsschnittstudie des Alterns*** (MLU-HistData), http://bolsa.uni-halle.de
- **bwFLA*** (ALU-RZ), http://eaas.uni-freiburg.de/
- **CAMO – Deutsche Akten im Zentralarchiv des russischen Verteidigungsministeriums (CAMO)*** (DHI Moskau), https://germandocsinrussia.org/
- **Central Description of the Collections of the Herder Institute*** (HI), https://www.herder-institut.de/holdings/?lang=en
- **Clio-online*** (HU), http://clio-online.de
- **ConedaKOR*** (IEG), https://github.com/coneda/kor
- **Corpus der Quellen zur Geschichte der Juden im spätmittelalterlichen Reich*** (AdW-Mainz, UT), http://medieval-ashkenaz.org/forschungsprojekt.html
- **DANTE*** (VZG), https://dante.gbv.de/
- DARIAH CAMPUS (DARIAH-EU), https://campus.dariah.eu/
- **DARIAH-DE Data Modelling Environement (DME)*** (DARIAH-DE, OFU), https://de.dariah.eu/dme
- **Das jüdische Hamburg*** (IGDJ), www.dasjuedischehamburg.de
- DataCite (TIB), https://www.tib.eu/de/publizieren-archivieren/doi-service
- **Daten-Eingabe-System (DES)*** (CompGen), http://des.genealogy.net/
- **DDB: Deutsche Digitale Bibliothek*** (FIZ, LABW, BSB, SBB-PK u.a.), https://www.deutsche-digitale-bibliothek.de/
- **Deutsche Auswandererbriefsammlung (DABS)*** (FB Gth, UT), http://www.auswandererbriefe.de
- **Deutsche Historische Bibliographie (DHB) | German Historical Bibliography (GHB)*** (BSB, DM, IFZ), https://www.historicum.net/dhb
- **Deutsches Museum Digital*** (DM), https://digital.deutsches-museum.de/
- **DIAMANT-Model*** (UT), http://fdm.uni-trier.de

- **digicultweb*** (digiCult), https://www.digicult-verbund.de/de/digicultweb
- **digiPress*** (BSB), https://digipress.digitale-sammlungen.de/
- **DIMAG*** (LABW, HLA), https://dimag-wiki.la-bw.de/xwiki/bin/Public/
- **DME (Data Modeling Environment)*** (OFU), https://de.dariah.eu/dme
- EAD (Encoded Archival Description) (LoC), https://www.loc.gov/ead/
- **eAQUA*** (UL-AG), http://www.eaqua.net/
- **EGO – Europäische Geschichte Online*** (IEG, UT), http://ieg-ego.eu/
- **EHRI-Portal*** (IFZ), https://www.ehri-project.eu/
- ENCHOS (European Network for the Comparative History of Population Geography and Occupational Structure 1500-1900) (University of Cambridge), https://www.campop.geog.cam.ac.uk/research/projects/internationaloccupations/enchpopgos/
- EOSC: European Open Science Cloud (EOSC), https://www.eosc-portal.eu/
- **Erlangen CRM*** (IGSD, DM, GNM), http://erlangen-crm.org/
- Europeana (Europeana Foundation), https://www.europeana.eu/
- **FactGrid*** (UE-FZG), https://database.factgrid.de/
- **FDM services*** (LRZ), https://www.lrz.de/forschung/projekte/forschung-daten/
- **FID Africa Studies*** (GU), https://www.ub.uni-frankfurt.de/projekte/afrikastudien.html
- **FID Book, library and information science*** (HAB, UL), https://katalog.fid-bbi.de/
- **FID Classics (Propylaeum)*** (BSB, HDU-UB), propylaeum.de
- **FID CrossAsia*** (SBB-PK), https://crossasia.org/
- **FID Educational Science and Educational Research*** (DIPF, BBF), https://www.fachportal-paedagogik.de/en
- **FID History*** (BSB, DM), historicum.net
- **FID Jewish Studies*** (GU), https://www.ub.uni-frankfurt.de/judaica3/home.html
- **FID Latin America*** (IAI), https://fid-lateinamerika.de/
- **FID Mid-East, North-Africa and Islam Studies*** (MLU), https://bibliothek.uni-halle.de/projekte/fid_nahost/
- **FID Musicology*** (BSB, SLUB), https://www.musiconn.de/
- **FID Russian, East and Southeast European Studies*** (BSB), https://www.osmikon.de/servicemenue/ueber-uns/fachinformationsdienst-ost-ostmittel-und-suedosteuropa
- **FID Social and Cultural Anthropology*** (HU), https://sozialundkulturanthropologie.fid-lizenzen.de
- **FoKO (Forschungsinfrastruktur Kunstdenkmäler in Ostmitteleuropa)*** (HI), https://foko-project.eu
- forschungsdaten.info (UK), https://www.forschungsdaten.info/
- forschungsdaten.org (DINI, GFZ), https://www.forschungsdaten.org/
- **FRET (Forschungsdatenrepositorium Trier)*** (UT), http://fret.uni-trier.de
- **FuD*** (UT), fud.uni-trier.de
- **GND*** (DNB), https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html
- **GND4C: GND for Cultural Data*** (DNB), https://www.dnb.de/EN/Professionell/ProjekteKooperationen/Projekte/GND4C/gnd4c.html
- **Gnomon Bibliographische Datenbank*** (UA, BSB), https://www.gbd.digital/metaopac/start.do?View=gnomon
- Google Dataset Search (Google), https://datasetsearch.research.google.com/

- **GOV: Geschichtliches Orts-Verzeichnis*** (CompGen), http://gov.genealogy.net/search/index
- **H-Soz-U-Kult*** (HU), hsozkult.de
- **Hamburg's Jewish cemeteries*** (IGDJ), http://www.jüdischer-friedhof-altona.de/datenbank.html
- **Handschriftenportal*** (SBB-PK, BSB, HAB, UL-UB), https://handschriftenportal.de/
- **HeFDI-RD-Repository*** (h_da u.a.), https://www.uni-marburg.de/de/hefdi/
- HISCO (Historical International Standard of Classification of Occupations) (IISG), https://iisg.amsterdam/en/data/data-websites/history-of-work
- **Historischer Atlas von Bayern*** (BSB, KBL u.a.), https://www.bayerische-landesbibliothek-online.de/hab
- **hypotheses Blog Portal*** (DHI Paris, OpenEdition), hypotheses.org
- **ICOM CIDOC WG Semantic Research Environments*** (DM, GNM, FAU), http://cidoc.mini.icom.museum/working-groups/semantic-research-environments/
- **IEG Maps*** (IEG), https://www.ieg-maps.uni-mainz.de/
- **IIIF (International Image Interoperability Framework)*** (BSB u.a.), https://iiif.io/
- **iRODS*** (ZDV), https://mogonwiki.zdv.uni-mainz.de/dokuwiki/start:fs_dm:archiving:irods
- **IStG Geodaten-Repositorium (Portal Städtegeschichte)*** (IStG), https://www.uni-muenster.de/Staedtegeschichte/portal/
- Journal of Digital History (JDH), https://journalofdigitalhistory.org
- **K10plus*** (VZG), https://wiki.k10plus.de/
- **Kalliope*** (SBB-PK u.a.), https://kalliope-verbund.info/
- **Kartenspeicher*** (VZG), http://kartenspeicher.gbv.de/index.php?id=3
- **Key Documents of German-Jewish History*** (IGDJ), jewish-history-online.net
- KldB2010: Klassifikation der Berufe 2010 (BfA), https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010/KldB2010-Nav.html
- **KONDA***: Kontinuierliches Qualitätsmanagement von dynamischen Forschungsdaten zu Objekten der materiellen Kultur unter Nutzung des LIDO-Standards (DDK), https://www.uni-marburg.de/de/fotomarburg/forschung/laufende/konda
- **KultSam*** (DM), https://www.deutsches-museum.de/forschung/forschungsbereiche/digitale-projekte/digitale-projekte/kultsam/
- **LAUDATIO*** (HU-CMS), http://www.laudatio-repository.org/
- **LEO-BW – Landeskundliches Informationssystem Baden-Württemberg (LABW),** https://www.leo-bw.de
- LIDO (Lightweight Information Describing Objects) (ICOM), http://www.lido-schema.org/
- MADOC (ULM), https://madoc.bib.uni-mannheim.de/
- **Map Collection of the Herder Institute*** (HI), https://www.herder-institut.de/go/bnu-3d3d1
- MARC: MAchine-Readable Cataloging (LoC), https://www.loc.gov/marc/
- METS/MODS: Metadata Encoding and Transmission Format (METS); Metadata Object Description Schema (MODS) (LoC), https://www.loc.gov/standards/mods/
- **Munich DigitiZation Center (MDZ)*** (BSB), https://www.digitale-sammlungen.de/
- **Museum4punkt0*** (DM), https://www.museum4punkt0.de/
- **OCR-D*** (SBB-PK), https://ocr-d.de/
- OER Commons (Institute for the Study of Knowledge Management in Education), https://www.oercommons.org/

- **OhdAB*** (MLU-HistData), https://www.geschichte.uni-halle.de/struktur/hist-data/ontologie/
- **Online-Findbuch der Dokumentesammlung des Herder-Instituts/Archiv zur Baltischen Geschichte*** (HI), https://www.herder-institut.de/actaproweb/index.xhtml
- **Oral-History.Digital*** (FU-CeDiS), https://www.cedis.fu-berlin.de/services/projektentwicklung/aktuell/ohd/index.html
- **OstData*** (BSB, HI u.a.), https://www.bsb-muenchen.de/ueber-uns/projekte/forschungsdatendienst-fuer-die-ost-ostmittel-und-suedosteuropaforschung-ostdata/
- **OstDok: Osteuropa-Dokumente Online*** (BSB u.a.), https://www.osmikon.de/servicemenue/ueber-uns/ueber-ostdok/
- **Picture Database on Jewish life*** (IGDJ), https://www.bilddatenbank-juedische-geschichte.de/
- **Presseausschnittarchiv des Herder-Instituts*** (HI), https://www.herder-institut.de/go/zL-c448f0
- programming historian (ProgHist Limited), https://programminghistorian.org/
- PST system of classifying occupations (UoC), https://www.campop.geog.cam.ac.uk/research/occupations/datasets/coding/
- **RADAR*** (FIZ), https://www.radar-service.eu/
- Ranke.2 (C2DH), https://ranke2.uni.lu/
- RDMO (AIP), rdmo.aip.de
- **Regesta Imperii*** (AdW-Mainz), http://www.regesta-imperii.de/startseite.html
- **Repositorium für Digitalisate*** (HLA), https://landesarchiv.hessen.de/dienststellen/digitales-archiv
- **RM.Net (Informationsnetzwerk zur Geschichte des Rhein-Maas-Raumes)*** (UT), rmnet.uni-trier.de
- **Sammlung Perthes Gotha*** (FB Gth), https://www.uni-erfurt.de/forschungscampus-gotha/campus-gotha/akteure/wissen-global/ihr-weg-zum-bestand
- segu Geschichte: selbstgesteuert-entwickelnder Geschichtsunterricht (segu), https://segu-geschichte.de/
- **share_it*** (MLU-HistData), https://opendata.uni-halle.de/
- **Standard Thesaurus for Economics*** (ZBW), https://zbw.eu/stw/version/latest/about.en.html
- TEI: Text Encoding Initiative (TEI), standard for the representation of texts in digital form (TEI Consortium), https://tei-c.org/
- Time Machine Europe (Time Machine Organisation), https://www.timemachine.eu/
- **Transcribo*** (UT), transcribo.org
- Transkribus (UI), transkribus.euc
- **URN-Service*** (DNB), https://www.dnb.de/DE/Professionell/Services/URN-Service/urn-service_node.html
- **ViDa*** (UT), http://vida.uni-trier.de
- **Virtuelles Kartenforum 2.0*** (SLUB), https://kartenforum.slub-dresden.de/
- **WissKI*** (IGSD, DM, GNM), https://www.drupal.org/project/wisski
- **xTree*** (digiCult), http://xtree-public.digicult-verbund.de/vocnet/
- **Zeitgeschichten Open*** (IFZ, BSB), https://open.ifz-muenchen.de/client/#/
- **ZfDG*** (HAB), http://www.zfdg.de/