



SYNTHEMA

D6.4

Project website

AUS



Funded by
the European Union

© SYNTHEMA 2022-2026

Revision v1.0

Work package	WP6
Task	T6.2
Due date	28-02-2023
Submission date	28-02-2023
Deliverable lead	AUSTRALO
Version	v1.0
Authors	Arantxa Echarte (AUS), Diana López (AUS)
Reviewers	Gustavo Hernández (UPM), Anna Rizzo (DW)

Abstract

The SYNTHEMA website (www.synthema.eu) includes key information about the project and aims to communicate to both professionals and wider audiences.

Keywords

About, team, structure, use cases, haematological data, health data, haematological diseases, federated computing networks, artificial intelligence techniques, synthetic data, virtual patients, rare disease, Sickle Cell Disease (SCD), Acute Myeloid Leukaemia (AML), privacy-by-design, ethics and data protection, clinical data anonymisation, European Health Data Space, secure multi-party computation

Document revision history

Version	Date	Description of change	Contributor(s)
v0.1	23-01-2023	1 st version of deliverable template	Arantxa Echarte (AUS), Diana López (AUS)
v0.2	28-02-2023	Deliverable content revision	Anna Rizzo (DW)
v0.3	28-02-2023	Deliverable content revision	Arantxa Echarte (AUS), Diana López (AUS)
V1.0	28-02-2023	Final version of the deliverable	Gustavo Hernández (UPM)

Disclaimer

The information, documentation and figures available in this deliverable are provided by the SYNTHEMA project's consortium under EC grant agreement **101095530** and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

Copyright notice

© SYNTHEMA 2022-2026

Document information

Nature of the deliverable*	DEC
Dissemination level	
PU Public, fully open. e.g., website	✓
CL Classified information as referred to in Commission Decision 2001/844/EC	
SEN Confidential to SYNTHEMA project and Commission Services	

* Deliverable types:

R: document, report (excluding periodic and final reports).

DEM: demonstrator, pilot, prototype, plan designs.

DEC: websites, patent filings, press and media actions, videos, etc.

OTHER: software, technical diagrams, etc.

Table of contents

1	Executive summary	6
2	Introduction to SYNTHEMA	7
3	Visual identity.....	7
4	Website.....	8
4.1	Site map.....	8
4.2	Main sections and pages.....	9
4.2.1	HOME	9
4.2.2	ABOUT	11
4.2.3	USE CASES	13
4.2.4	TEAM	15
4.3	Additional sections	17
4.3.1	NEWS.....	17
4.3.2	CONTACT	17
4.3.3	PUBLICATIONS	17
5	Conclusions.....	18

List of figures

Figure 1.	SYNTHEMA – Site map	8
Figure 2.	SYNTHEMA – HOME page.....	10
Figure 3.	SYNTHEMA – ABOUT page.....	12
Figure 4.	SYNTHEMA – USE CASES.....	14
Figure 5.	SYNTHEMA – TEAM.....	16

Acronyms and definitions

AI	Artificial Intelligence
AML	Acute Myeloid Leukaemia
DoA	Description of Action
DP	Differential privacy
EHDS	European Health Data Space
ENROL	European Rare Blood Disorders Platform
ERDRI	European Rare Disease Registry Infrastructure
EU RD platform	European Platform on Rare Disease Registration
FL	Federated learning
KPI	Key Performance Indicators
GDPR	General Data Protection Regulation
MDS	Myelodysplastic syndrome
RHD	Rare haematological diseases
SCD	Sickle Cell Disease
SDG	Synthetic Data Generation
SME	Small and Medium Enterprise
SMPC	Secure multi-party computation
WP	Work Package

1 Executive summary

In order to systematically increase the public visibility of SYNTHEMA and subsequently also its outcomes, a project website has been created (www.synthema.eu). The public project website enables the general public and relevant stakeholders to be informed about the project objectives and approach, its news and results and it provides the contact information of the project partners.

Other communication tools (such as branded templates, flyers, online publishing, etc.) as well as dissemination activities will be described in detail in **D6.1 – Impact master plan**, which will be released in M6.

The website will be the primary asset for promoting project activities and results to all target audiences, providing comprehensive information about its objectives, team, events, pointing at outside related resources. The project will consider eco-friendly hosting providers with servers in the EU ([A2Hosting](#)). Moreover, the website follows the EU recommendation regarding usability and accessibility, and it includes the logo of the European Commission.

The website sections are:

- **HOME:** the Home section includes a general description, mission and challenges, objectives and concept of the project.
- **ABOUT:** the About section includes information about the European Health Data Space (EHDS), the main goal and the key drives for the platform.
- **USE CASES:** the Use Cases section includes an introduction to Haematological diseases and the two specific Use Cases: *Sickle cell disease* (SCD) and *Acute myeloid leukaemia* (AML).
- **TEAM:** this section includes an introduction to [ERN-EuroBloodNet](#) and [ENROL](#) networks, and also details about the project partners, including a link to their websites for further information.
- **NEWS:** a collection of news pieces and updates related to the project.
- **CONTACT:** this section provides information on how to contact the project team and request further information about the project.

Through the website we aim to comply with specific project objectives and KPIs, like for example, communication, dissemination, multimedia (teaser video, interviews, short clips) distribution, website visitors (300 unique visitors/month on average) and non-scientific publications. Additionally, open access to the raw datasets via the individual repositories will be ensured through the SYNTHEMA community profile on specialised public access databases like [Zenodo](#), [DataHub.io](#), [Frictionless Data](#) and [Open-AIRE](#), Portal of Medical Data Models, among others. Hyperlinks to all access points will be given on the SYNTHEMA website.

All partners will contribute to the website contents. However, the day-to-day management will be AUSTRALO's responsibility.



2 Introduction to SYNTHEMA

SYNTHEMA aims to establish a privacy-preserving, cross-border hub to develop and validate innovative *Artificial intelligence* (AI) models for clinical data anonymisation and synthetic data generation in rare hematological diseases.

Haematological diseases are a large group of disorders resulting from abnormalities in blood cells, lymphoid organs and coagulation factors. Rare hematological diseases inherently suffer from data scarcity and fragmentation, but there is also the dilemma of data privacy and protection. Can we talk about full anonymity when the risk of re-identification is so high? How can we generate meaningful synthetic data to circumvent the lack of quality data to train AI models on?

The overarching ambition of SYNTHEMA is to increase the number of existing samples in this disease space (with a focus on two highly representative use cases: SCD and AML), thus fighting off the critical issues of data scarcity and fragmentation and pushing the boundaries of patient-centric, GDPR-compliant research.

SYNTHEMA enjoys the support, resources and active participation of [ERN-EuroBloodNet](#), as [the European Reference Network on rare haematological diseases \(RHDs\)](#) concentrating 103 highly specialised multidisciplinary healthcare teams in 24 Member States. Moreover, the [European Rare Blood Disorders Platform \(ENROL\)](#), conceived in the core of [ERN-EuroBloodNet](#) in line with the EC strategy for Rare Diseases as an umbrella for new and existing RHDs registries, directly contributes to SYNTHEMA on the promotion of interoperability standards of the EU RD platform to tackle the scarcity and fragmentation of data and widen the basis for GDPR-compliant research in RHDs.

3 Visual identity

In order to reach a broad external as well as internal audience, a project visual identity has been created. The main goal is to establish and convey a coherent image and brand recognition leading to an optimal presentation and recognition of the project.

The visual identity uses a set of graphic elements to easily identify the project. The overall aim is to create a coherent and highly recognisable image of the project to support communication and dissemination measures, such as publications and all kinds of written as well as visual communication about on-going and completed research activities.

Additionally, social media banners, common templates for written deliverables (MS Word, see format of this report), minutes of meeting and presentations (MS PowerPoint) have been created. In pursuance of a consistent brand identity, the design of the templates follows the same style of the logos.

Further information on the project visual identity regarding its function and implementation in other dissemination tools will be described in detail in D6.1 (Impact master plan).

4 Website

The project public website is a key instrument to maximise visibility of the project and the main entry point to showcase its main ideas and approach, news, findings and results, introducing visitors to the project rationale and educating them on its concept.

4.1 Site map

A complete sitemap is presented below. The website structure and navigation menus will be expanded and fine-tuned as the project progresses and there is more content hosted online.

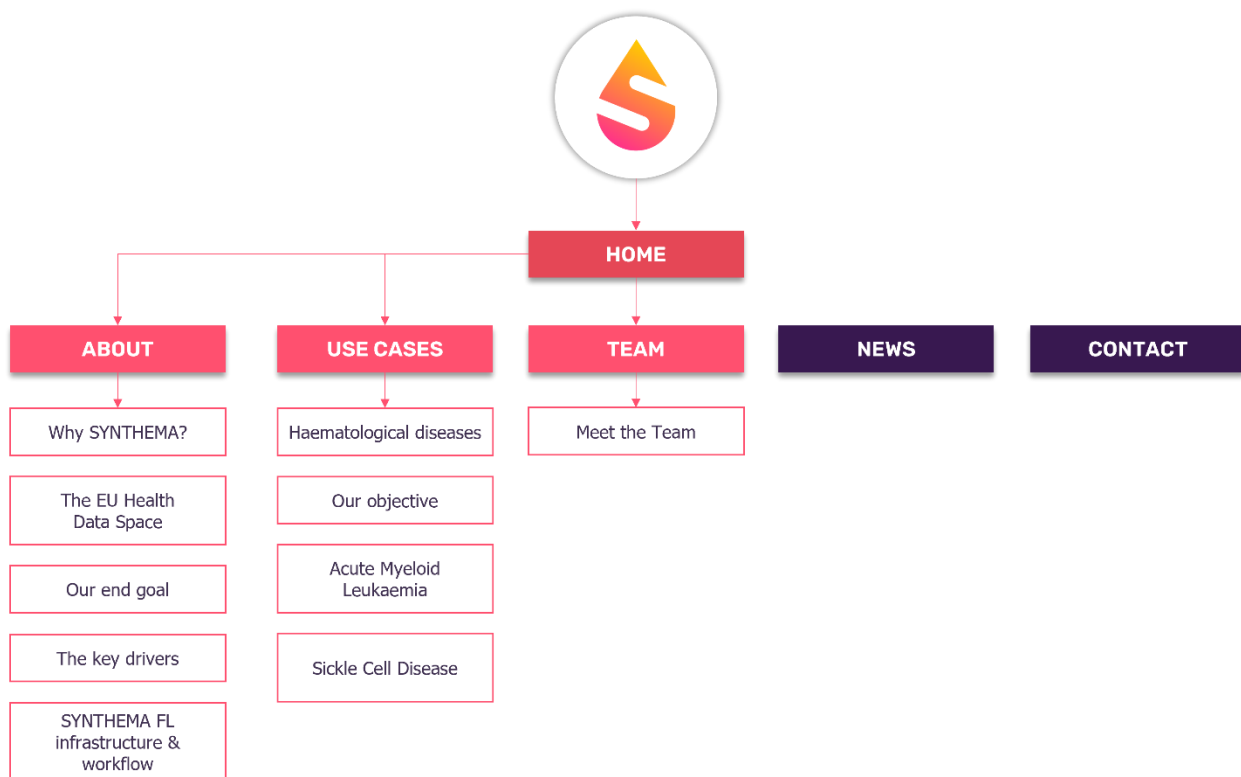
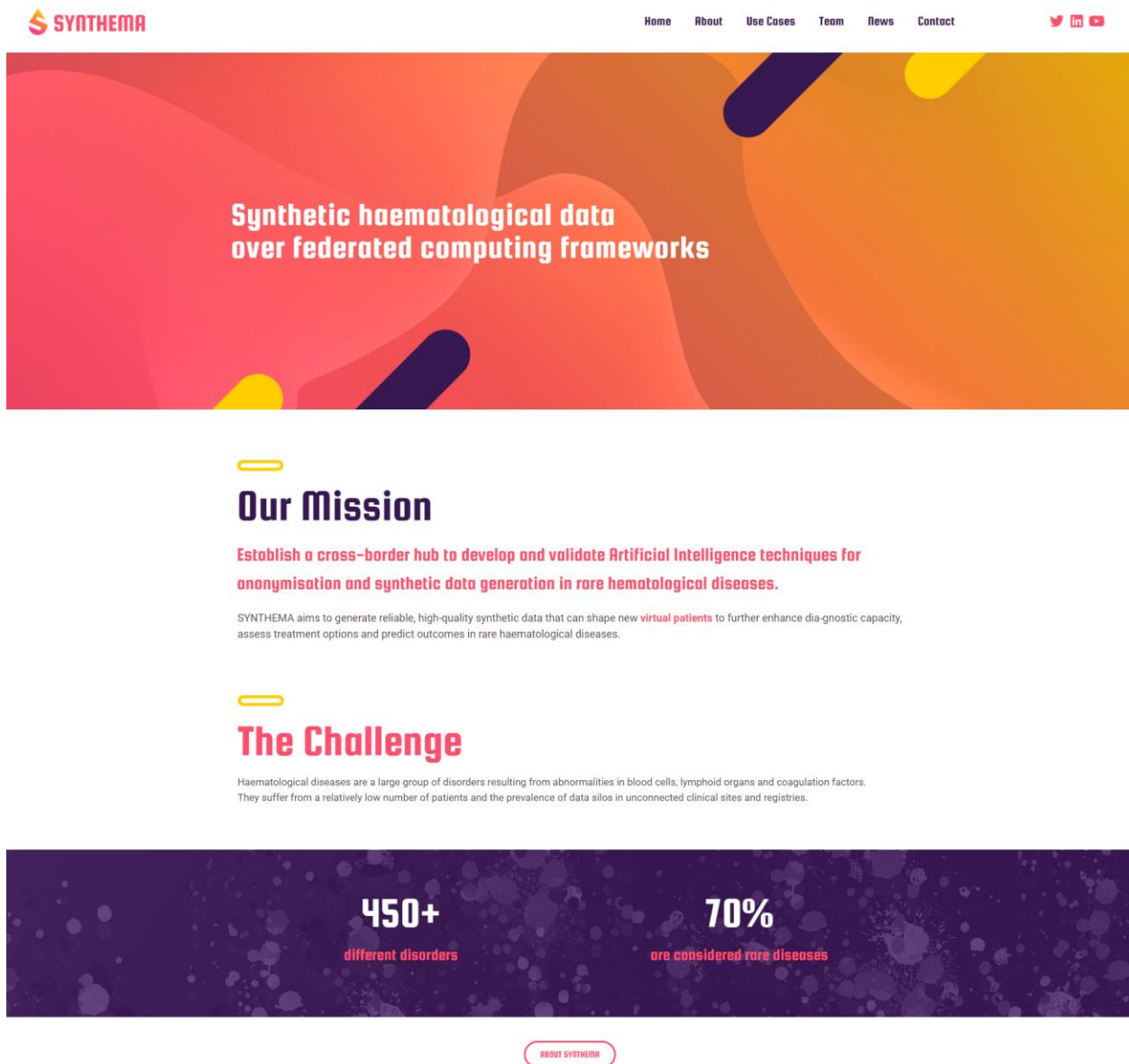


Figure 1. SYNTHEMA – Site map

4.2 Main sections and pages

4.2.1 HOME

The website landing page. This section includes our “main headlines” for the project – it is a straightforward, visually attractive, to-the-point overview of the project vision, goals and our key message. The homepage has been designed to grab the user’s attention and guide them effectively to any other section on the website to keep them engaged.



The screenshot shows the SYNTHEMA website homepage. At the top left is the SYNTHEMA logo. To the right is a navigation menu with links for Home, About, Use Cases, Team, News, and Contact. Further right are social media icons for Twitter, LinkedIn, and YouTube. The main content area features a large hero section with a colorful abstract background (red, orange, yellow, and dark blue) and the headline: "Synthetic haematological data over federated computing frameworks". Below this is a section titled "Our Mission" with a yellow underline, followed by a red sub-headline: "Establish a cross-border hub to develop and validate Artificial Intelligence techniques for anonymisation and synthetic data generation in rare hematological diseases." Below this is a paragraph of text: "SYNTHEMA aims to generate reliable, high-quality synthetic data that can shape new virtual patients to further enhance diagnostic capacity, assess treatment options and predict outcomes in rare haematological diseases." Another section titled "The Challenge" with a yellow underline follows, with a red sub-headline: "Haematological diseases are a large group of disorders resulting from abnormalities in blood cells, lymphoid organs and coagulation factors. They suffer from a relatively low number of patients and the prevalence of data silos in unconnected clinical sites and registries." At the bottom of the page is a dark blue banner with white and red text: "450+ different disorders" and "70% are considered rare diseases". Below the banner is a red button with white text: "ABOUT SYNTHEMA".

Our work in the rare disease space

We are devoted to expanding the landscape of personalized medicine in rare haematological diseases. We aim to increase the number of existing samples in the rare haematological disease space. SYNTHEMA focuses on two highly representative use cases: **Sickle Cell Disease (SCD)** and **Acute Myeloid Leukaemia (AML)**.



Virtual patients

Novel methods to generate synthetic multimodal clinical, omics & imaging data



Privacy-by-design

Built-in privacy through the combination of federated learning (FL), secure multi-party computation (SMPC) and differential privacy (DP)



Ethics and data protection

Ensuring value-sensitive design for AI development and GDPR compliance for data collection and processing



Collaboration is key

Embracing open science practices and multidisciplinary training in rare hematological

[BROWSE OUR USE CASES](#)

A multidisciplinary team

Funded by the European Union under Horizon Europe, this initiative embodies the pan-European collaboration of top experts, selected for their acknowledged excellence and complementarity bringing knowledge, expertise and state-of-the-art clinical, computational, business, ethical and legal background while minimizing overlaps.

4

Years

16

Partners

10

Countries

[MEET THE TEAM](#)

Subscribe to our newsletter

[MEET THE TEAM](#)



[Privacy Policy](#) | [Terms Of Use](#)

FOLLOW US!

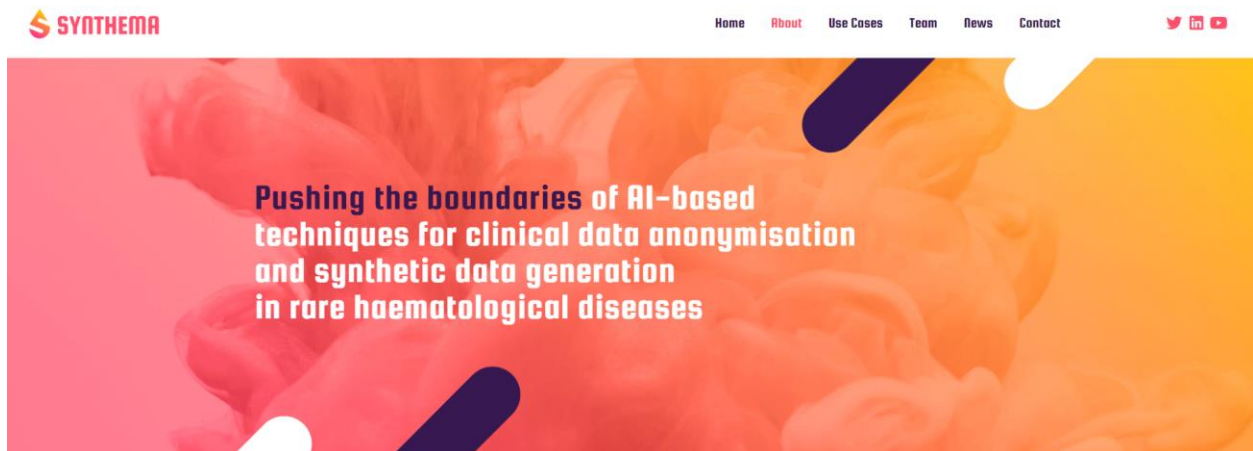


©SYNTHEMA2023. All rights reserved

Figure 2. SYNTHEMA – HOME page

4.2.2 ABOUT

The About page offers a more in-depth look into the project innovations and methodology. Here, we have showcased more detailed information about SYNTHEMA context, targets, key drivers and overall technical framework.



Why SYNTHEMA?

Under a global lens, the impact of haematological diseases is staggering.

Unfortunately, as is common for rare diseases, both the scarcity and fragmentation of available data prevent researchers from reaching the critical mass required for pushing forward basic and clinical research, and heavily impacts health authorities' capabilities for effective health planning.

Moreover, their underrepresentation in coding systems highlights the complexity in tracing patient pathways within healthcare systems, thus limiting the long-term sustainability of existing and new patient registries established at national and European level.

30m+

rare disease patients in the EU

15.3B€

are considered rare diseases

This is exactly why **precision medicine is key**: when we shift our focus to the individual, all diseases become unique.

[DIVE INTO OUR USE CASES](#)

The European Health Data Space

The **European Health Data Space (EHDS)** is one of the central building blocks of a strong **European Health Union**. It aims to create a strong legal framework for the use of **health data** for research, innovation, public health, policy-making and regulatory purposes. Under strict conditions, researchers, innovators, public institutions or industry will have access to large amounts of high-quality health data, crucial to develop life-saving treatments, vaccines or medical devices and ensuring better access to healthcare and more resilient health systems.

SYNTHEMA shares the vision of the EHDS to create a consistent, trustworthy and efficient framework to use health data to push the boundaries of clinical research, while ensuring full compliance with the EU's high data protection standards.

Our end goal

SYNTHEMA aims to establish a cross-border health data hub for rare haematological diseases: a space to develop and validate innovative AI-based techniques for clinical data anonymisation and synthetic data generation (SDG). The ultimate ambition is to address the issues around data scarcity and fragmentation to effectively widen the basis for meaningful, GDPR-compliant research in this disease space.

The key drivers

The SYNTHEMA platform will be based on a privacy-preserving federated learning (FL) network, equipped with secure multi-party computation (SMPC) protocols and differential privacy (DP), connecting health data and academic research centres, industries and SMEs to advance translational and clinical research and care working in rare haematological diseases.



Synthetic data generation

To generate virtual patients that keep patterns and features of real-world data



Federated Learning

To facilitate the collaborative training of AI models with no sharing of raw data



Secure multi-party computation

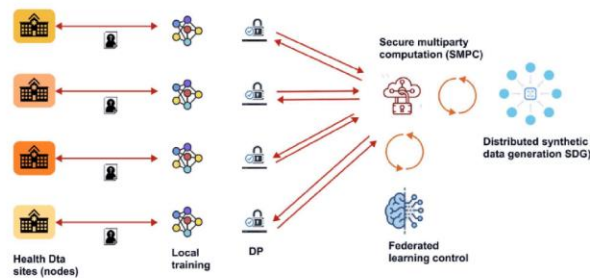
For lower risk, privacy-preserving model aggregation in FL schemes



Differential privacy

To set strict boundaries on the disclosure of private data each clinical site is allowed

SYNTHEMA FL infrastructure architecture and workflow



Subscribe to our newsletter

MEET THE TEAM



Funded by the European Union



Privacy Policy | Terms Of Use

FOLLOW US!



©SYNTHEMA2023. All rights reserved

Figure 3. SYNTHEMA – ABOUT page

4.2.3 USE CASES

This section provides an overview of the different disease use cases planned within the project. In our case, the spotlight is on AML and SCD. For both, this page contextualizes the disease and outlines the project goals, focus and partners involved in the effort.

The screenshot shows a website page with a red-to-orange gradient background featuring a pattern of translucent spheres. At the top left is the SYNTHEMA logo, and at the top right is a navigation menu with links for Home, About, Use Cases, Team, News, and Contact, along with social media icons for Twitter, LinkedIn, and YouTube. The main heading reads "Haematological diseases at the forefront of rare disease research". Below this, a section titled "Haematological diseases" includes a paragraph: "Haematological diseases are generally divided into oncological (haematological malignancies, i.e., lymphomas, myelomas, leukaemias) and non-oncological HDs (i.e., hemoglobinopathies, haemolytic anaemias, coagulopathies). Even though over 74% of them are considered rare, the overall number of HD affected patients worldwide is significant: in 2020 alone, haematological malignancies accounted for about 5% of total cancer cases worldwide." A second section titled "Our objective" features a progress bar with four steps: Provide, Develop, Ensure, and Contribute. Below the progress bar, it states: "Provide novel methods and capabilities to generate synthetic multimodal (clinical, omics and imaging) data for rare haematological diseases with a validated clinical result". The final section, "Our use cases", is set against a dark purple background with a faint microscope graphic. It lists two use cases: "Acute Myeloid Leukaemia" and "Sickle Cell Disease", each with a "LEARN MORE" button.

Acute Myeloid Leukaemia

The disease

Acute Myeloid Leukemia (AML) is a type of blood cancer that starts from young white blood cells in the bone marrow. The bone marrow is the soft inner part of the bones, where new blood cells are made.

The bone marrow produces white blood cells called granulocytes or monocytes too quickly because they grow and divide too fast. These abnormal cells build up in the blood and bone marrow and can eventually spread to other parts of the body including the lymph nodes and the spleen.

It is the most common type of acute leukemia in adults. AML can get worse quickly if it is not treated, however, treatments work very well for most people.

Validation

For AML, validation scenarios will assess performance of synthetic data in determining overall survival and treatment response prediction based on clinical-genomic profiles.

Cohort

The AML use case will include 2,500+ patients with "de novo" AML (2016 World Health Organisation/WHO classification criteria) and 2,000+ patients with myelodysplastic syndrome (MDS), a pre-leukemic condition that can evolve into AML and allows us to study the transition between the chronic and acute disease phase.



Sickle Cell Disease

The disease

Sickle Cell Disease (SCD) is a group of hereditary red blood cell disorders. It is a rare, chronic and life-threatening disease. In patients with SCD, red blood cells become C-shaped in resemblance to a sickle, the farming tool the disease is named after.

Sickle cells die early and tend to clog the blood flow when going through small blood vessels, so patients usually suffer from low red blood cell counts, infections, acute chest syndrome and strokes.

Treatment for sickle cell disease is limited. However, following improvements in preventative treatment, many complications associated with sickle cell disease can be avoided or delayed.

Validation

For SCD, validation scenarios will test the reliability of synthetic data in regards to genomic variants/disease phenotypes association and MRI feature-based prediction of brain vascular events (SCD).

Cohort

The SCD use case will include 1,000+ patients with genetic diagnosis for SCD disorder, including paediatric (1+ year-olds) and adult patients (exclusion criteria: transplanted patients, 1 year-old or younger).

Subscribe to our newsletter [HIDE THE FORM](#)

Funded by the European Union

[Privacy Policy](#) | [Terms Of Use](#)

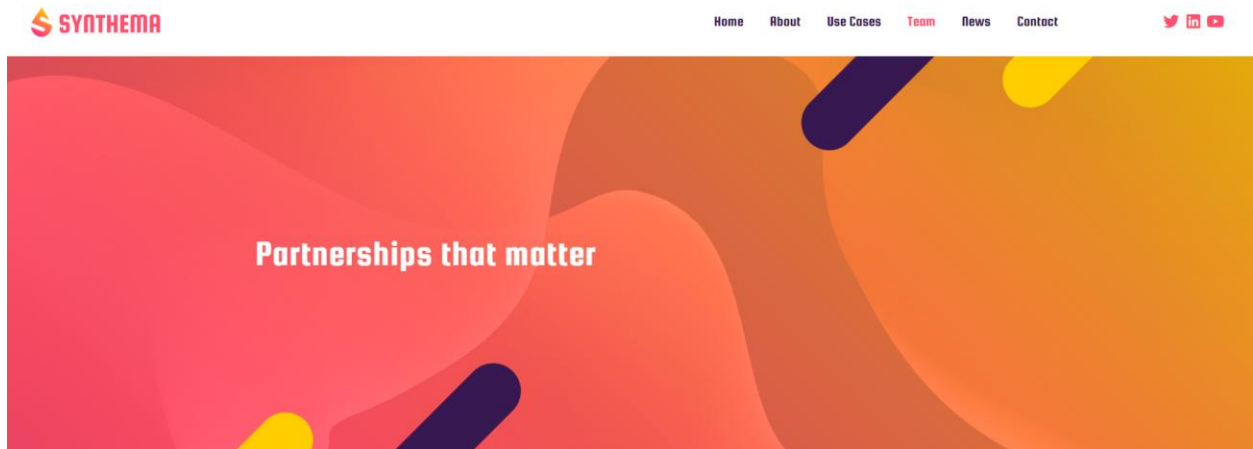
FOLLOW US!

©SYNTHEMA2023. All rights reserved

Figure 4. SYNTHEMA – USE CASES page

4.2.4 TEAM

A sneak peek of the partners in the consortium, including general information and contact details on each specific partner.



16 partners in 10 countries

16 partners in 10 countries will join forces to push the boundaries of synthetic data in rare haematological diseases. SYNTHEMA also enjoys the support, resources and active participation of **ERN-EuroBloodNet**, as the European Reference Network on rare haematological diseases (RHDs) concentrating 103 highly specialised multidisciplinary healthcare teams in 24 Member States.

Our platform will contribute to existing data registries such as the European Rare Blood Disorders Platform (ENROL), the European Platform on Rare Disease Registration (EU RD Platform) and the European Rare Disease Registry Infrastructure (ERDRI) with data standards, pipelines and shareable data assets.

SYNTHEMA will connect health data centres, academic research centres, industries and SMEs to advance transnational and clinical research and care in RHDs. All in all, ERN-EuroBloodNet and ENROL constitute the perfect environment for SYNTHEMA and the creation of a cross-border health data hub for RHDs, as the space to develop and validate innovative AI-based techniques for clinical data anonymisation and synthetic data generation.



Meet the Team

Haematological diseases are a large group of disorders resulting from abnormalities in blood cells, lymphoid organs and coagulation factors. They suffer from a relatively low number of patients and the prevalence of data silos in unconnected clinical sites and registries.



Subscribe to our newsletter [MEET THE TEAM](#)

Funded by the European Union

Privacy Policy | Terms Of Use

FOLLOW US!

©SYNTHEMA2023. All rights reserved

Figure 5. SYNTHEMA – TEAM page

4.3 Additional sections

4.3.1 NEWS

The news section will provide any news and events related to the project, such as consortium meetings, workshops, participation at conferences, and it will also provide access to dissemination material related to the project (logo, flyer, newsletters, press release, etc.).

4.3.2 CONTACT

A simple contact form that users can fill in to reach out to the project office for collaboration opportunities or further information.

4.3.3 PUBLICATIONS

As the project progresses, all public documentation and resources (e.g. public deliverables, Open Access publications...) will also be accessible through the project official website.

5 Conclusions

The aim of the present deliverable **D6.4 – Project website** is to give a detailed overview on the project visual identity created to be used throughout the project. By developing a professional joint image and appearance a sound basis for further dissemination and exploitation activities has been therefore set already in the beginning of the project.

The website (www.synthema.eu) will go officially live on 28-02-23, and it will be updated on a weekly basis for the first three months –due to the initial activity– and bi-weekly thereafter. The look of the SYNTHEMA website is currently as illustrated above and it follows the project graphic identity, however the design is still in progress and some additional improvements are planned.

