Health Data Spaces

Interoperable, Scalable and Sustainable Health Data Spaces in Europe

Gaia-X health ecosystem position paper: a path towards fully federated and interoperable health data spaces at scale.

17 November 2026

Editor

Bert Verdonck (Luxembourg National Data Service).

Primary authors

Anja Burmann (Fraunhofer ISST; sphin-X e.V.), **Cecile Badiola-Lagardere** (Heva-data), **Hubert Tardieu** (Gaia-X), **Jorge Rodríguez Graña** (NTT Data), **Pierre Gronlier** (Gaia-X), **Sukumar Munshi** (European Health Data Alliance), **Zia Alborzi** (NTT Data).

Contributors

Alberto Borrego Diez (NTT Data), **Alberto Curatolo** (Gaia-X), **Anupam Choudhary** (Luxembourg National Data Service), **Florian Lauf** (Fraunhofer ISST), **Julian Cervinoiglesia** (NTT Data), **Michael Kamp** (Lamarr Institute, TU Dortmund), **Michael Mossal** (NTT Data), **Ulrich Ahle** (Gaia-X).

Table of Contents

Executive summary 3

Introduction 4

Country status 6

- Country status: France 6
- Country status: Germany 8
- Country status: Spain 10

Stakeholder expectations 14

How to make EHDS successful? 15

How to make health data spaces trustable? 17

How to make health data spaces interoperable? 19

How to make health data spaces sustainable? 23

Call to action 24

Executive summary

Problem statement. Europe's health data landscape remains extremely fragmented across institutions, between countries and various infrastructures. Cross-border data flows are limited, access to high-quality data is uneven, and successful digital health and AI pilots rarely scale. This fragmentation restricts better care and efficient research and innovation, with an untapped potential of a population of more than 450 million citizens.

Vision. This paper defines a European ecosystem of federated, interoperable and trustworthy health data spaces. This ecosystem enables secure data use within and across borders while preserving institutional control, national responsibilities and European sovereignty. It supports affordable healthcare, high-quality research, public-health policy and competitive health innovation, including vertical and agentic AI.

Approach. The paper positions the European Health Data Space Regulation as the governance backbone. Gaia-X provides operational mechanisms for identity, trust, policy enforcement and verifiable compliance, while DSSC2 supplies reusable building blocks and a components toolbox. National and sectoral health data spaces form the implementation layer. Together, these elements translate regulation into scalable processes for data discovery, access, interoperability, secure processing and auditability.

Contributions. Experiences from France, Germany and Spain demonstrate different national pathways towards federated health-data infrastructures. The paper identifies reusable governance, trust and technical patterns, including interoperable onboarding, common metadata, semantic standards, exchangeable secure processing environments, verifiable credentials and end-to-end interoperability testing. It establishes a practical pathway towards trusted cross-border implementations at scale.

Key outcomes. The paper identifies interoperability, trust, sovereignty and sustainability as the foundations of a functioning European health-data ecosystem. It calls for a shift from isolated, time-limited pilots to durable infrastructure, common operating models and reusable services. This approach connects health data spaces with Europe's research and AI capabilities and makes high-value use cases repeatable across institutions, disease areas and Member States. The paper serves as an invitation to continue the dialogue with other contributors and other initiatives that share the same objectives.

Call to action. The authors request that health data space initiatives shall:

1. Scale health data spaces across borders.
2. Build a minimum common interoperability and trust baseline.
3. Create durable funding and operating models.
4. Connect health data spaces to the European AI, innovation and research agenda.

Introduction

Affordable and sustainable healthcare is at stake globally, due to the demographic evolution, the increasing burden of chronic disease, and the workforce shortage in the health sector. Europe has an extremely fragmented health information landscape, and, in many regions, antiquated health information software installed. Cross-border data flows are rare, to deliver care to cross-border workers and migrating citizens as well as for aggregating health data for research and clinical studies. Digital health use cases and pilots do not scale (yet). Europe has yet to unlock the full potential of its data at the scale of its 450 million citizens, to advance both healthcare and research.

Vertical AI in the health domain ("Health AI") depends on trustworthy, federated data spaces as a foundation. For sensitive health data, the current data aggregation methods as used for Generative AI are clearly not possible: sensitive health data shall not be aggregated exclusively and exploited by large commercial players. This is also essential to prepare for Agentic AI applications: when agents work on sensitive health data, automated access and trust management will become a necessity.

The European Health Data Space regulation (EHDS) is poised to solve a major part of these problems. The regulation is since long awaited, will still take many years to deploy, and is a necessary but not sufficient data sharing framework. The regulation addresses the rights of citizens, imposes obligations for health data holders and for health information system providers, and defines the infrastructures for cross-border data sharing.

EHDS addresses:

- the availability of patient data across the EU region for better care (EHDS I),
- the availability of clinical and other data for better research (EHDS II),
- the issues with inferior interoperability and lock-in effects related to vendor specific technologies.

To make EHDS fully operational, multiple challenges and implementation gaps remain to be solved:

- there are no standard recipes for implementation within Member States, as EHDS is limited to cross-border and cross-European exchanges;
- there is no joint view and no joint adoption of best interoperability practices across Member States, apart from the European Health Record Exchange Format (EHRxF)[1] which is great for patient care but insufficient for many clinical study needs;
- there remain ambiguities between the regulation, the implementation guidelines and the (yet to be finalised) implementation acts;
- there is a lack of support for data related operations, such as for the reuse of curated datasets and the persistence of data mapping and lineage;
- there are no specific solutions to prepare data for vertical AI;

[1] *For EHRxF see https://ehr-exchange-format.eu/ and https://www.hl7europe.org/standards/eehrxf/*

- the time frames of EHDS implementation and local data space initiatives are out of sync, creating issues with integration and consistent use of the various infrastructures.

The purpose of this paper is to craft a joint perspective of leading health data space initiatives on how to build interoperable and sustainable health data spaces within the context of the EHDS. It builds upon an earlier positioning paper from the Gaia-X health ecosystem, published in 2021 by a group of Gaia-X members active in the health domain[1]. And it is not an end product: the paper serves as an invitation to continue the dialogue with other contributors and other initiatives that share the same objectives.

The paper provides guidance on in-country as well as cross-border use cases and on scaling within and beyond single Member States.

The paper targets the health leadership stakeholders, such as policymakers (DG Health, DG Connect, Member State Ministries of Health, Research, Digitalisation and the Economy), the health industry, healthcare providers, patient representatives and research organisations.

[1] *Previous Gaia-X health data space white paper, part of Gaia-X position paper of August 2021 (page 99 onwards): https://gaia-x.eu/wp-content/uploads/files/2021-08/Gaia-X_DSBC_PositionPaper.pdf*

Country status

We decided to focus on three European Member States with some of the largest populations[2]: France (69M), Germany (83M) and Spain (50M). Together they represent 45% of the total 451M population of the European Union. They also have very differently organised health systems, with various levels of federalisation, making them interesting to compare.

Country status: France

The French health data ecosystem is already highly structured and considered one of the most advanced in Europe. At its core is the Health Data Hub (HDH)[3], created in 2019 as a public platform to facilitate secure access to health data for research, innovation, and public policy purposes. The HDH relies on the Système National des Données de Santé (SNDS), which covers nearly the entire French population and includes claims data, hospital data, prescriptions, and mortality records, making it one of the most comprehensive health databases worldwide.

The governance of health data in France is strongly state-driven, with the Ministry of Health and the Delegation for Digital Health (DNS) defining strategies and coordinating implementation, particularly in relation to European regulations. The Health Data Hub acts as the operational arm, while the CNIL[4] ensures compliance with data protection and GDPR rules, and institutions such as CNAM[5] remain key data providers and stewards.

At the European level, the adoption of the European Health Data Space (EHDS) regulation in March 2025 marks a major turning point. The Health Data Hub's experience in managing access requests and coordinating data sharing is relevant to the future role of health data access bodies, whose national designation is still in progress.

To prepare for EHDS implementation, several national and European initiatives are underway. In particular, the French HealthData@EU project, co-funded by the European Commission, aims to improve the visibility, interoperability, and reuse of French health data through the development of standardised metadata catalogues and improved data quality processes.

The FeData project is a major French initiative aiming to build a federated national infrastructure for health data, connecting all University Hospitals (CHU) across the country in a unified but non-centralised model. It is designed to allow researchers and industry players to access multicentric data through a single entry point, while keeping data hosted locally within each hospital's data warehouse.

This approach is based on a federated architecture, where data are not pooled centrally but remain within the original institutions, ensuring better governance, sovereignty and scalability. The project

---

[2] EU population data: https://doi.org/10.2908/DEMO_GIND
[3] HDH = Health Data Hub, https://www.health-data-hub.fr/ or PDS = Platforme des Données de Santé
[4] CNIL = Commission Nationale de l'Informatique et des Libertés, https://cnil.fr/en
[5] CNAM = Caisse nationale de l'assurance maladie, https://www.ameli.fr/ and https://www.assurance-maladie.ameli.fr/etudes-et-donnees/en-savoir-plus-snds

currently involves approximately 32 CHU, which together represent the backbone of clinical data production in France.

FeData is strategically positioned as a key enabler of clinical research and artificial intelligence, by facilitating faster cohort identification, improving access to real-world data, and enabling large-scale multicentric studies. It also contributes to building a sovereign and trusted data ecosystem, aligned with both national priorities and European ambitions. The project is supported by a public-private consortium, notably involving CHU, Docaposte/HEVA (La Poste Santé & Autonomie), and ecosystem partners, combining operational, technological, and research capabilities.

From an industry perspective, there is strong and growing demand for access to high-quality, large-scale health data, particularly for real-world evidence generation, AI development, and clinical research optimisation. Industry players expect faster and more predictable access processes, clearer governance frameworks, and interoperable infrastructures capable of supporting large-scale use cases. In this context, initiatives such as FeData illustrate the shift towards federated data models, which are increasingly seen as a pragmatic compromise between scalability, sovereignty, and trust.

However, despite significant progress, several structural challenges remain in France. First, access to health data remains complex and time-consuming, due to the multiplicity of stakeholders, including CNIL authorisations, ethical approvals, and negotiations with data holders. Second, interoperability issues persist, as hospital information systems, data standards and their application remain heterogeneous, requiring major efforts in harmonisation and data quality improvement.

Third, data sovereignty has been a sensitive issue given the historical hosting of the Health Data Hub on Microsoft Azure. A transition to a European hosting provider was decided in 2026 and is currently being implemented.

Fourth, public trust and societal acceptance remain critical, as health data are highly sensitive. Although the use of such data is strictly regulated and legally validated, ongoing debates highlight the need for transparency, strong governance, and clear communication to citizens.

Finally, a key challenge lies in the operational implementation of EHDS, as there is still a gap between the regulatory ambition at European level and the actual readiness of national infrastructures, governance mechanisms, and technical capabilities, including the articulation between federated infrastructures and the future national access bodies.

In conclusion, France benefits from strong structural assets in health data, notably the health insurance data SNDS[6] and the Health Data Hub, and is progressing towards more scalable, interoperable and sovereign models. In this trajectory, initiatives such as FeData illustrate a federated approach to large-scale and trusted access to and analysis of health data, designed to be compatible with the European Health Data Space.

[6] *SNDS = Le système national des données de santé, https://www.assurance-maladie.ameli.fr/etudes-et-donnees/en-savoir-plus-snds*

Country status: Germany

Germany combines nationwide digital-health infrastructure with decentralised research networks and a growing portfolio of Gaia-X-aligned data-space initiatives. Its opportunity is to connect these assets through a clear application of the EHDS role models, binding interoperability and sustainable operating incentives.

Current situation

Germany's health data landscape is broad but fragmented across insurance, care settings, 16 federal states, registries, university hospitals, research networks and private platforms. Core assets include the telematics infrastructure (TI), electronic patient record (ePA), Health Data Lab (FDZ Gesundheit), Data Access and Coordination Office (DACO), Medical Informatics Initiative (MII) with its Data Integration Centres (DIZ), the German Research Data Portal for Health (FDPG), NUM, NFDI4Health, cancer registries and genomDE.

The ePA as German Electronic Health Record (EHR) has been broadly rolled out as an opt-out infrastructure since 2025. By April 2026, gematik reported almost 73 million records, more than 100 million stored documents and weekly access by up to 93,000 medical and dental practices[7]. This demonstrates national reach, while semantic completeness, cross-sector adoption and systematic reuse remain work in progress.

Care	ePA and TI provide the national backbone; cross-border primary-use obligations phase in from 2029.[11]
Research	MII/DIZ, FDPG, NUM, NFDI4Health, registries and genomDE form a strong federated network.
Secondary use	DACO coordinates access; FDZ and prospective domain-specific HDABs provide domain-level functions.
Industry	Pharma, MedTech, health-IT providers and SMEs complement the public core with data space and analytics services.

National policies, strategies and driving ministries

The Federal Ministry of Health (BMG) leads health-sector digitalisation, EHDS implementation and the legal framework for ePA, FDZ and DACO. 'GEMEINSAM DIGITAL 2026' links healthcare delivery, systematic data generation, AI and European connectivity[8]. BMFTR supports research infrastructure and transfer, BMWE data spaces and industrial scaling, and BMDS digital sovereignty and horizontal digital policy. BSI, BfDI, state data protection authorities, the federal states and self-governing organisations remain essential implementation actors.

[7] *gematik: One year of the ePA for all - key figures, 15 April 2026*

[8] *BMG: GEMEINSAM DIGITAL 2026*

HDAB assignment and implementation plan

The Health Data Use Act (GDNG) established the Data Access and Coordination Office (DACO) at the Federal Institute for Drugs and Medical Devices (BfArM) as the central coordination point for health-data access. The July 2026 GeDIG draft develops this into a decentralised EHDS model with one coordinating body and domain-specific health data access bodies (dHDABs). As legislation and implementing ordinances are not complete, these assignments remain a target model rather than final law [9] [10].

- DACO: central intake and routing, multi-domain coordination, complaints and connection to HealthData@EU; HDP4Germany[11] develops the application, metadata, secure-processing and gateway components.
- FDZ Gesundheit: operational access and analysis for statutory health-insurance claims data; proposed as the domain-specific access body for this domain.
- FDPG / MII-DIZ: established workflow for harmonised clinical data and bio samples. It is an important operational component, but not automatically a legally designated dHDAB.
- Further domains: the GeDIG explanatory memorandum identifies the Centre for Medical Registries for registry data and NUM for university medical data.

Ongoing health data-space initiatives

Public and research infrastructure. DACO/HDP4Germany provide the coordination and European connectivity layer; FDZ enables regulated analysis of claims data; and FDPG, MII-DIZ and NUM provide a federated university-medicine pathway. FDPG plus is extending feasibility, application, contracting and transparency functions. In March 2026, FDPG documented more than 21 million hospital stays, over two billion laboratory values and more than 250 million medication data points[12].

Industry and Gaia-X ecosystem. Data4Good from the European Health Data Alliance (EHDA) develops federated AI models, participatory governance and a viable operating approach[13]. sphin-X[14] brings together more than 30 organisations in over 20 working groups and use cases, including digital supply chains and EHR-to-EDC (Electronic Data Capture). HEALTH-X dataLOFT tested a citizen-centred data-space architecture and concluded as a funded project in 2025, with follow-on activities continuing orchestrated by the European Health Data Alliance[15].

European and cross-border links. The IDERHA[16] project is developing a disease-agnostic federated data space for research, regulatory use and Health Technology Assessment (HTA) through 2028.

[9] *BMG: Health Data Use Act (GDNG)*
[10] *BMG: Federal Cabinet draft of GeDIG, July 2026*
[11] *HDP4Germany: https://www.bfarm.de/EN/BfArM/EU-and-International/EU-Projects/EU4Health/HDP4Germany/_node.html*
[12] *MII: FDPG plus - functions of the German Research Data Portal for Health*
[13] *EHDA: Data4Good*
[14] *sphin-X e.V.: health data ecosystem and use cases*
[15] *HEALTH-X dataLOFT: project page and final event*
[16] *IDERHA: federated European health data space*

Luxembourg's Dataspace4Health[17] project demonstrates Gaia-X identities, HealthDCAT-AP metadata, policy and contract negotiation, and EDC-based data transfer. These initiatives offer reusable cross-border patterns, but do not replace national HDAB decisions or EHDS permits.

Industry position

Pharmaceutical and MedTech companies expect better access to high-quality data for real-world evidence, clinical development, product safety, health technology assessment, personalised medicine and AI validation. Platform, cloud and connector providers focus on secure processing environments, federated analytics, identity, data quality and auditability. SMEs and start-ups could benefit from discoverable datasets and standardised access routes, but face proportionally high compliance, preparation and integration costs.

Across industry, the main requirements are predictable application timelines, transparent fees, reusable compliance evidence, meaningful data-quality descriptions and protection of trade secrets. Concerns centre on regulatory overlap between the EHDS, GDPR, AI Act, MDR/IVDR, Data Act and NIS2; weak incentives for data holders; and the risk that incompatible access portals and secure processing environments create new silos.

Remaining challenges and priorities

Legal and organisational clarity: define responsibilities, escalation, supervision and liability across DACO, dHDABs, data holders, intermediaries and secure-processing operators.

Interoperability: make FHIR profiles, terminology, HealthDCAT-AP metadata, provenance and data-quality metrics binding and testable across sectors.

Funding and sustainability: move from time-limited pilots to operating models with stable financing, transparent fees, shared storage and compute infrastructure, open interfaces and incentives for data holders.

Adoption and trust: integrate solutions into provider workflows, reduce administrative burden, protect commercial interests and demonstrate tangible value to citizens and patients.

Potential contribution towards 2030: Germany can provide a reusable blueprint for large, fragmented health systems: a public access core for legal certainty, federated data holders for institutional autonomy, and Gaia-X/DSSC2 components as the trust and exchange layer. The next value step is to validate this model through a small portfolio of high-impact cross-border use cases and a sustainable funding pathway, including an IPCEI-like or EU-co-financed second phase.

Country status: Spain

Spain is building a multi-level and increasingly federated health-data ecosystem reflecting the decentralised Spanish National Health System, where Autonomous Communities retain major responsibility for healthcare delivery and health data.

[17] *Gaia-X: Dataspace4Health - August 2026 status*

The central national initiative is the Espacio Nacional de Datos de Salud (ENDS)[18], launched on 29 January 2026 and led by the Ministry for Digital and Civil Service Transformation, the Ministry of Health and the Autonomous Communities. Spain has allocated approximately EUR 70 million to national ENDS infrastructure, plus EUR 28 million for technological adaptation of regional health information systems[19].

ENDS enables secure secondary use of health data for research, innovation, healthcare improvement and evidence-based policy. Its value lies in a federated model suited to distributed responsibilities and data assets: instead of replacing regional capabilities, ENDS connects them through common governance, protocols and interoperability mechanisms while preserving distributed control of data across hybrid regional, on-premises and cloud environments.

The ENDS catalogue uses HealthDCAT-AP[20] for dataset discovery and interoperability with the European Health Data Space (EHDS). Secure Processing Environments enable controlled secondary use: authorised users work remotely with approved datasets, raw data cannot be extracted, and only approved aggregated results may leave after security review. ENDS therefore provides a national coordination and interoperability layer for secondary health-data use under the EHDS.

ENDS also sits within broader Spanish data-space investment. Sectoral programmes led to 2024–2025 funded demonstrators, reusable technologies and health-related initiatives, while "Kit Espacios de Datos" supports onboarding of public and private organisations into operational national and regional data spaces. The result is a growing health-sector portfolio spanning healthcare, personalised medicine, biomedical research, One Health[21] analytics, AI and federated sharing.

EHDS implementation and governance

Spain is already implementing operational components required by Regulation (EU) 2025/327 while the national statutory framework matures. The HDAB-ES project, led by the Ministry of Health and co-financed through EU4Health, runs from 2023 to 2027 and supports Health Data Access Body capabilities, including access requests, the national dataset catalogue, secure processing and integration with HealthData@EU.

Implementation is coordinated through the SNS Data Office (ODSNS) within a governance model combining common policies and standards with decentralized implementation by the Autonomous Communities and their health services. Madrid illustrates the regional path: in 2026, it launched work on Health Data Access Body capabilities covering access requests, permits, secure processing and dataset cataloguing, with the catalogue designed to federate with ENDS and HealthData@EU.

The legal framework is still evolving. A proposed Digital Health Law is intended to adapt Spain's framework to the EHDS Regulation, while Article 18 of Royal Decree 415/2026 provides that ENDS will serve as a secure processing environment for Spain's Health Data Access Bodies. Spain is

[18] https://ends.gob.es/datos-recursos/infraestructura-del-ends
[19] https://www.lamoncloa.gob.es/lang/en/gobierno/news/paginas/2026/20260129-national-health-data-space.aspx
[20] https://ends.gob.es/datos-recursos/healthdcat-ap
[21] https://www.cqssalud.com/en/the-spanish-government-promotes-the-one-health-approach-to-protect-human-animal-and-environmental-health/

therefore already operationalising significant EHDS capabilities while completing the legal and institutional framework.

Regional and specialised initiatives

The regional landscape demonstrates how this federated model is materialising in practice.

In Madrid, HealthData@MAD-R&I is developing a federated environment for secure secondary use in research and innovation, using interoperability approaches including OMOP. SHARE – SpanisH cAnceR data spacE, completed in May 2026, connected clinical information on more than 30,000 lung-cancer patients across 197 Spanish hospitals while keeping personal data within hospitals and sharing only analytical results, showing that distributed clinical datasets can support large-scale research without centralising sensitive patient-level data.

In Galicia, OneHealth DataSpace involved 36 entities and 43 active use cases by July 2026 across human, animal and environmental health, combining secure sharing with analytics, AI and high-performance computing, while work is progressing towards future interoperability and integration with ENDS. In Asturias, Cuélebre provides regional health-data capabilities for secondary use, research, analytics and AI, while FINBA Data Trust adds federated data-space capabilities for secure inter-organisational collaboration.

FINBA[22] is especially relevant to Gaia-X because its publicly documented architecture aligns with Gaia-X and IDSA principles, uses open-source EDC connectors, and includes a Gaia-X Digital Clearing House provided by ARSYS for conformity verification. It is a concrete Spanish example of Gaia-X-oriented trust and conformity mechanisms in a health data space.

Private-sector participation

Private-sector participation is also becoming tangible. The Espacio de Datos de la Sanidad Privada (EDSP), led by Fundación IDIS and launched in July 2026 with 14 participating organisations, provides a decentralised environment for secondary use within the private-healthcare ecosystem. Participants retain data custody, publish metadata using HealthDCAT-AP and use controlled processing mechanisms where only data required for an approved purpose enters a Secure Processing Environment, and only aggregated outputs may be extracted.

EDSP is designed with future EHDS connectivity in mind. Private participation is also emerging directly within ENDS: in April 2026, SEOGA, operated by Ulex Soluciones, became the first private participant to formalise adhesion to ENDS and publish an initial dataset in the national catalogue.

Scaling this participation will require predictable access processes, data quality and discoverability, clear governance, reusable onboarding and sustainable operating models.

Challenges and Gaia-X relevance

Spain's main next-stage challenge is interoperability across independently governed environments. Technical connectivity alone will not be sufficient: effective federation requires interoperability across organisational, legal, semantic, technical and operational layers, including trust, access procedures, metadata, standards, data quality, secure processing and operational

[22] *https://fdt.finba.es/proyecto/*

processes. Sustainability is equally important: demonstrators and publicly funded projects need paths toward durable services and repeatable operating models.

In this context, Gaia-X should be viewed as complementary to ENDS and the EHDS, not as an alternative infrastructure. The EHDS establishes the European regulatory and governance framework for health-data access and secondary use; ENDS provides national coordination and infrastructure capabilities; and Gaia-X can contribute reusable mechanisms for trust, identity and verifiable compliance across independently governed ecosystems. FINBA already illustrates this complementary role.

The resulting opportunity is to develop Spain's growing portfolio of national, regional, specialised and private initiatives into a coherent, trusted and interoperable health-data ecosystem capable of supporting secondary-use cases at national and European scale.

Stakeholder expectations

From the country profiles above, the major stakeholder expectations can be summarised as in the following table:

Stakeholder	Main value	Main concern	Mandatory practice
Citizens and patients	Better care, faster diagnosis, more effective research	Privacy, misuse, lack of transparency	Clear information, lawful use, visible safeguards
Care providers	Better clinical information, reduced duplication, better outcomes	Extra workload, liability, poor system integration, lack of funding	Seamless workflows, trust in data quality, interoperable tools
Researchers	Access to larger and more diverse datasets	Slow permits, fragmented access, poor metadata	Discoverability, predictable authorisation, efficient access, secure analysis
Industry	Reliable access for R&D, validation, regulatory and HTA use	Unclear timelines, fragmented governance, confidentiality risks	Standardised access routes, efficient access, secure processing, legal certainty
Public authorities and HDABs	Better public health, supervision and policy capacity	Compliance burden, unclear responsibilities, lack of operational tools, lack of sustained funding	Traceability, auditability, common processes and accountability

The TEHDAS2 project published[23] a good overview of the allowed purposes for the secondary use of health data. They describe valuable and relevant use cases:

- Public interest – protect against health threats and ensure quality and safety of healthcare.
- Policymaking and regulatory activities: Support public bodies – help government and EU health institutions.
- (Official) Statistics – create health-related statistics.
- Education – for teaching in health and care.
- Scientific Research related to health – for innovation of new health products and testing, testing of algorithms, and for public health research
- Personalised healthcare – provide tailored healthcare based on others' data.

We observe that the various health data space initiatives across Europe are pursuing very common requirements and use cases. Furthermore, we observe that the solution designs, architectures, pilot and production implementations have many areas of overlap and that many aspects are reinvented, piloted and tested many times. Therefore, we conclude that current investments are scattered and that synergies are underutilised. We expect that much more speed and much larger scale can be created by harmonising solutions and by clubbing investments.

How to make EHDS successful?

The European Health Data Space regulation (EHDS) is the first extensive regulatory framework within the context of the Common European Data Spaces, fully dedicated to the health data ecosystem. According to the European Commission, the EHDS will[24]:

1. empower individuals to access, control and share their electronic health data across borders for healthcare delivery (primary use of data) à enable the availability of patient data across the EU to deliver better care;
    a. enable seamless health data access across borders
    b. enable secure sharing of medical histories supporting better diagnosis and treatment decisions
    c. reduce unnecessary duplicated medical tests
    d. support the transition from paper-based to digital records
    e. support telemedicine and personalised treatments and cross-border collaboration for improved healthcare delivery

[23] *TEHDAS2 Deliverable D6.2 – Guideline for data users on good application and access practice: https://tehdas.eu/wp-content/uploads/2025/10/guideline-for-data-users-on-good-application-and-access-practice.pdf*
[24] *EHDS on EC website: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en*

2. enable the secure and trustworthy reuse of health data for research, innovation, policy-making, and regulatory activities (secondary use of data) à enable the availability of health data across the EU for better research;
    a. advance digital healthcare
    b. facilitate data-driven research and innovation, enhance efficiency and affordability in medical advancements
    c. empower patients with greater control over their health data (opt-out provisions)
3. foster a single market for electronic health record (EHR) systems, supporting both primary and secondary use à overcome the data silos between vendor specific technology

With the EHDS now on its path to implementation, multiple challenges and implementation gaps remain to be solved. A recent EOSC health data task force report[25] lists the gaps in hosting, discoverability and accessibility, in interoperability and data quality, and in data processing. The EIT Health think tank report identified gaps in governance, capacity/skills and resources/funding[26].

In our view, these are the major remaining challenges:

- no standard recipes for implementation within Member States
- no joint view and adoption of best "interoperability" practices across Member States; major ontology issues remain
- ambiguities between regulation, implementation guidelines and yet to be finalised implementation act *(this might be a temporary issue)*
- no solutions for data related operations such as the aggregation of curated datasets
- no specific solutions to prepare data for vertical AI
- time frames of EHDS implementation and local initiatives are out of sync, creating issues with integration and consistent use

The Digital Health Commons Forum has worked on several of these challenges, published about the systemic failures in digital health[27] and worked on some guiding documents to address the challenges above[28]. They set out a practical case for an open, sustainable digital health ecosystem. It combines foundational framing with operational playbooks, investment logic and governance guidance, and it explains what is broken, what must change, and how organisations can act without reinventing standards, platforms or institutions.

Based on this diverse prior work, the following sections will suggest solutions and recommendations to address the challenges mentioned.

[25] *Gaps report from EOSC health data task force, 27 March 2026, https://doi.org/10.5281/zenodo.19254262*
[26] *EIT Health think tank report, April 2024, https://eithealth.eu/think-tank-topic/implementing-the-european-health-data-space/*
[27] *Digital Health Commons Forum publication on the digital health landscape: https://dhcf.eu/publications/the-digital-health-landscape*
[28] *Digital Health Commons Forum publications: https://dhcf.eu/publications/*

The intended solution architecture is multi-tiered: first a catalogue/metadata and discoverability layer, then potentially aggregate queries or federated analytics, and only for an authorised use case: access to patient-level data, preferably or exclusively in a secure processing environment. The main features of this architecture are now described in the following chapters on trust, interoperability and sustainability. Aspects of digital sovereignty are mentioned as a cross-cutting concern. Please note that trust and interoperability do not imply centralisation or direct access to patient-level data.

How to make health data spaces trustable?

Trust is a precondition for scaling health data spaces across borders. In practice, cross-border data sharing cannot rely on bilateral due diligence, manual checks and one-off contracts alone. Hospitals, registries, Health Data Access Bodies, researchers, companies and infrastructure providers must be able to establish sufficient confidence in one another's identity, authority, compliance status and behaviour. Trustability therefore means the ability of a health data space to provide verifiable and reusable evidence that its participants, services and transactions operate under agreed legal, organisational and technical conditions.

The EHDS provides the essential legal and governance framework for primary and secondary use of health data in Europe. It defines rights, obligations and institutional roles, and it establishes the broader European framework within which national implementations must operate. However, regulation alone does not make cross-border trust operational. A hospital in one Member State still needs practical means to verify whether a researcher, service provider, secure-processing environment or automated service from another Member State is authorised, compliant and accountable. This is why a trust framework is needed in addition to regulation: it connects legal and governance requirements to everyday operating processes.

What needs to be trusted in a health data space goes beyond the dataset itself.

First, **organisations** must be trusted: hospitals, research organisations, Health Data Access Bodies, service providers and companies need to be identifiable and properly onboarded.

Second, **natural persons and roles** must be trusted: for example, whether a person is authorised to act as a researcher, data steward, clinician or security officer.

Third, **services and infrastructure** must be trusted: connectors, catalogues, identity services, secure processing environments and analytics tools must demonstrate that they meet agreed requirements.

Fourth, **data and metadata** must be trusted: provenance, quality indicators, access conditions and permitted uses must be clear and reliable.

Finally, **transactions and outputs** must be trusted: requests, approvals, workload execution and output release must be logged, auditable and, where necessary, reversible.

For this reason, European health data spaces should converge on a minimum common set of trust guardrails. These should include verifiable identity for organisations, persons and machine

services; reusable onboarding supported by common credentials and attestations; authorisation linked to role and purpose; enforceable policies covering access, use, retention and output release; continuous compliance evidence; tamper-evident logging and auditability; secure processing and controlled output handling; and the ability to suspend or revoke permissions when risks, roles or circumstances change. Such guardrails should not force identical national governance systems. Rather, they should provide a common baseline that enables trust across independently governed health data spaces.

This is where Gaia-X and DSSC2 become practically relevant. EHDS defines the legal and institutional objectives; Gaia-X-style trust frameworks provide operational mechanisms to establish identity, verify claims, express policies and automate conformity checks across independent participants; DSSC2 complements this with reusable guidance and building blocks across governance, legal, business, technical and interoperability dimensions. Together, these frameworks help translate regulatory requirements into interoperable trust mechanisms that can be reused across national and sectoral health data spaces.

A particularly important concept is interoperable onboarding. Today, organisations often face repeated onboarding, repeated compliance checks and repeated bilateral negotiations when joining multiple data-sharing environments. This does not scale. The objective should therefore be “onboard once, reuse evidence many times, authorise locally where required.” Participants should not need to repeat the same audit, onboarding and compliance checks in every health data space they join.

A scalable approach is to issue key compliance and qualification evidence as machine-readable, verifiable credentials that can be checked automatically and cryptographically by counterparties. This does not replace human oversight, legal review or audit; it makes them faster, more repeatable and more portable across trusted environments.

Common identity, organisational credentials, conformance evidence and compliance claims should be reusable across multiple health data spaces, while decisions on data access, legal basis and health-specific permissions remain with the competent data holders, HDABs and related authorities. In this way, interoperable onboarding reduces friction without undermining national and institutional responsibilities.

Technology exists to support these capabilities: they are known as “automate compliance” in the regulatory technology domain. Providers issue attestations as machine-readable verifiable credentials that counterparties check cryptographically. Participants can build on interoperable, already operational trust frameworks rather than creating another isolated mechanism. The Gaia-X Trust Framework, born from Member States' digital policies and built by European organisations, shows how policy requirements become machine-readable rules, verifiable credentials and automated conformity checks.

Automation does not replace human control: audits and explainability remain essential; automated checks make them faster and repeatable, not optional. "Minimum common trust guardrails" should mean a common machine-readable set of identification, compliance and qualification requirements every health data space recognises (further developed in the Interoperability section).

Trustability needs to evolve for a world of increasing automation and agentic AI. In the future, software agents may help discover datasets, prepare requests, deploy approved workloads and support analysis. Yet such agents must never become unaccountable actors. Every autonomous service should operate under a clearly identified principal, within explicit and limited permissions, with complete logging, traceability and revocation mechanisms. Automation can accelerate approved workflows, but it must not bypass legal requirements, human accountability or data-holder sovereignty. Health data spaces should therefore extend their trust frameworks to cover not only organisations and people, but also machine services and AI agents.

The Gaia-X Trust Framework covers both regulatory compliance (natural and business entities) and, increasingly, policy compliance in human-to-agent and agent-to-agent interactions. Health data spaces will connect not only organisations but AI agents acting on their behalf - discovering datasets, negotiating access, executing agreements - each action requiring a trust decision at machine speed.

- Verifiable delegation for agents: agents must be bound to their organisation's verifiable credentials, with a cryptographically verifiable delegation chain - the only scalable way to control what an autonomous service may do.
- Machine-readable policy: express rules so agent actions can be checked before or at execution ("compliance by construction"), extending policy compliance from entity-level to interaction-level.

Europe needs a minimum interoperable trust baseline for health data spaces. This baseline should enable participants to demonstrate identity, authority, compliance and accountable operation across borders, while allowing Member States and data holders to retain legitimate governance responsibilities. Trustability is therefore not an optional technical add-on. It is the condition that allows EHDS to move from regulatory intent to trustworthy, scalable and reusable health-data operations in practice. Gaia-X-aligned trust mechanisms and DSSC2 implementation guidance provide a credible path to achieve this.

How to make health data spaces interoperable?

Europe’s health-data landscape is populated by a wide range of initiatives: national health-data infrastructures, electronic health record systems, disease registries, research platforms, hospital data warehouses, biobanks, data catalogues, secure processing environments and public-private data-sharing projects. These initiatives differ in purpose, scale, governance, maturity and technical architecture.

Interoperability across health data spaces thus means the conditions under which relevant infrastructures and projects can connect, interoperate, evolve or converge. This is consistent with the European Health Data Space (EHDS). EHDS establishes the legal rights, obligations and governance framework for primary and secondary use of health data. MyHealth@EU and HealthData@EU provide essential European connectivity mechanisms. However, regulation and connectivity alone do not guarantee that data, services and organisations can work together in

practice. Interoperability requires implementable agreements and evidence that these agreements work across systems and borders.

This requires an interoperability framework towards a coherent European health-data ecosystem. Governance models, data custody arrangements and national specificities can differ, especially once further domain regulations come into effect. A common set of interoperability requirements must still ensure practical processes across initiatives for discovery, access, trust, semantics, secure processing and auditability. Interoperability should therefore be treated as a strategic design criterion for national implementation, European funding and the further development of health-data infrastructures. The practical goal is an evolving and interoperable landscape of health-data capabilities. This landscape should support the EHDS while enabling Member States, healthcare providers, research organisations and industry to build on existing investments, reuse common components and scale proven cross-border use cases.

Health data spaces can only interoperate effectively when several layers are addressed.

- Legal and policy interoperability requires a shared and operational understanding of rights, legal bases, permitted purposes, access conditions and safeguards under the EHDS, GDPR and national law, including rules for intellectual property, trade secrets and output checking.
- Organisational and governance interoperability defines the roles and responsibilities of Health Data Access Bodies, data holders, data users, secure-processing operators and intermediaries, as well as common procedures for onboarding, authorisation, auditing, liability, escalation and mutual recognition.
- Semantic interoperability ensures that data can be understood and assessed consistently through common data models, implementation profiles, terminology services and governed mappings between international and national terminologies, accompanied by provenance and data-quality metadata.
- Operational interoperability makes data access repeatable in practice, covering discovery, feasibility assessment, permit workflows, data preparation, secure processing, output release, monitoring and incident management.
- Technical interoperability provides the secure machine-to-machine capabilities required to implement these processes, including standardised APIs, identity and access management, connectors, encryption, secure processing environments, logging and portable workloads.
- Finally, trust and sovereignty are cross-cutting requirements: participants, services and transactions must be identifiable, compliant with applicable rules, transparent, traceable and auditable, while data holders retain control over data location, access, permitted use and revocation.

A pragmatic approach from standards to interoperable implementation should focus on a limited set of high-value, reusable interoperability assets:

- Common metadata and discoverability. All datasets, data services, secure processing environments and relevant applications should publish machine-readable metadata. HealthDCAT-AP provides a strong basis for health-data catalogue interoperability,

complemented by information on data holder, purpose, access restrictions, provenance, quality, applicable standards and available processing environments.

- Common data and exchange profiles. Health data spaces should converge on mature standards where appropriate: HL7 FHIR and IHE profiles for health-information exchange, OpenEHR for longitudinal patient records, DICOM for medical imaging, OMOP for observational research and analytics, and GA4GH standards for genomics. The relevant choice depends on the use case; no single model will cover all healthcare, research and public-health needs.
- Terminology and semantic translation services. Semantic interoperability requires terminology binding, version management and transparent mappings. Member States should avoid treating national coding systems as barriers to European reuse. Where replacement is neither realistic nor desirable, health data spaces should provide governed, versioned and testable translation services between national and international terminologies.
- Data quality, provenance and fitness-for-purpose. A dataset must be assessable before it can be responsibly reused. Health data spaces should publish common quality indicators, such as completeness, timeliness, conformance, plausibility, representativeness and known limitations. Provenance must show where data originated, which transformations were applied and under which governance conditions it may be used.
- Interoperable access and secure processing workflows. Secure Processing Environments support interoperable application, authorisation, workload deployment, logging, output checking and audit processes. This makes it possible to bring approved analysis to data where data transfer is restricted or unnecessary.
- Authorisation-to-execution interoperability. Cross-border interoperability must connect the authorisation decision to technical enforcement. The authorised user, dataset, purpose, conditions, validity period and processing environment should be consistently translated into enforceable access and usage policies; data space mechanisms complement, not replace, the permit. Execution should then produce interoperable evidence of what was performed and released, creating a traceable permit-to-policy-to-execution-to-evidence chain across heterogeneous implementations.
- Make interoperability testable: Interoperability must be demonstrable, measurable and continuously maintained. Health data spaces should therefore adopt an interoperability assurance framework combining conformance testing, operational testing and governance evidence.

At minimum, this framework could assess whether a participant can:

- publish and discover metadata through a federated catalogue;
- authenticate organisations and authorised individuals across borders;
- validate relevant credentials, roles and compliance claims;
- submit, receive and process a standardised data-access request;
- interpret agreed semantic profiles and terminology bindings;

- demonstrate data quality and provenance information;
- enforce access conditions within a secure processing workflow;
- create tamper-evident logs and support audit;
- release only authorised, checked outputs;
- and revoke, suspend or update access when permissions, policies or circumstances change.

Testing should include end-to-end cross-border "test journeys" based on realistic use cases, for example a multi-county rare-disease cohort study, federated oncology image analysis, cross-border patient-summary access, or post-market safety monitoring. These exercises involve the full chain: data holders, HDABs, researchers, secure-processing operators, catalogue services and identity providers.

A European health-data-space test suite could define minimum conformance profiles and maturity levels. Certification or labelling should not be limited to infrastructure providers; it should also cover relevant services, connectors, metadata implementations, secure processing environments and organisational operating procedures. Reference implementations, sandbox environments and reusable onboarding packages are particularly important to lower entry barriers for hospitals, research organisations, SMEs and regional initiatives.

Zooming in on the interoperability of trust, onboarding is where trust becomes operational: proving who you are (identification), that you comply (compliance), and that your data and services qualify (qualification). Cross-border care and research should not force hospitals, research institutes or digital health providers to repeat this in each data space. Trust guardrails must be defined legal-first, then made executable down the stack:

- **Legal interoperability:** align applicable rules (EHDS, GDPR, Data Act, national health law) so legal entitlement in one health data space implies recognition in the others.
- **Organisational interoperability:** agree on issuer accreditation, trust framework governance and mutual recognition, so participants face one organisation-level conformity assessment, not one per data space. Gaia-X Trust Framework provides the verification machinery.
- **Semantic interoperability:** shared machine-readable ontologies and schemas for claims, so a credential issued in one data space is understood identically in another. This requires harmonising ontologies - in preparation and in parallel to EHDS - including automated translation between currently deployed standards.
- **Technical interoperability:** the standards that carry and verify claims - Gaia-X on W3C standards (decentralised identifiers, verifiable credentials), reaching partners beyond the EU for cross-border identification and mutual recognition.

Modern trust framework technologies, such as the Gaia-X Trust Framework, anchor trust in open standards, automated policy engines and machine-verifiable credentials, establishing it peer-to-peer through cryptographic verification of credentials from accredited issuers. Institutional mandates define the requirements; standards-based verification executes them.

In practice, decentralised identifiers and verifiable credentials (GDPR compliance, certifications, authorisations) flow down this stack: defined legally, governed organisationally, understood

semantically, verified technically - once, for every health data space. This is also where **metrics, testability and certification** come in: automated conformity checks make requirements measurable and repeatable across health data space initiatives.

How to make health data spaces sustainable?

Sustainability in health data spaces should be understood in three dimensions: societal sustainability, operational sustainability and economic sustainability. Societal sustainability means that citizens, patients and institutions continue to see public value in participation. Operational sustainability means that services remain usable, governed and maintained beyond project funding. Economic sustainability means that the costs of onboarding, compliance, curation, secure processing and interoperability are covered through durable financing and viable operating models, rather than through repeated short-term pilots.

We note that the (academic) literature on the economy of health data spaces is extremely limited to non-existent. Stakeholders are used to public funding to create the legal, technical and operational health data sharing solutions. When funding runs out, initiatives stop, for the right or for the wrong reasons. The broader health economic impact of efficient data sharing remains poorly understood. We call for increased socio-economic research on these matters.

For operational sustainability, the governance framework and governing authority are crucial. The Digital Health Commons Forum has documented best practices in designing the "custodian" role[29] for managing sustainable governance in health data ecosystems.

AI changes the sustainability equation for health data spaces. A data space creates lasting value when the same governed data assets and access mechanisms can support repeated high-value use cases, and AI is one of the strongest multipliers. Training, testing and validating health AI require diverse, high-quality and representative data that is currently fragmented across hospitals, registries, research infrastructures and countries. Rather than repeatedly copying these data into new central platforms, federated approaches can enable approved data and computation to meet under explicit access conditions, provenance and accountability. Where sensitive data should remain under institutional control, approved analytical or AI workloads can instead be brought to the data through secure processing or other distributed execution patterns. This makes investments in interoperability, data quality and trust reusable across multiple AI use cases instead of repeatedly funding isolated pilots.

Europe should therefore deliberately connect health data spaces with its emerging AI infrastructure. The EHDS provides the regulatory framework for legitimate secondary use, AI Factories provide AI-optimised compute capacity and expertise, and health data spaces can provide the governed domain-data ecosystem that makes these capabilities useful for healthcare and research. But AI-ready must mean more than making datasets available. Data should be sufficiently documented to assess fitness for purpose, including provenance, quality, representativeness and known limitations; be represented through interoperable semantic models

[29] *Custodians as described in the digital health commons: https://dhcf.eu/publications/the-role-of-the-custodians*

and standards; carry clear conditions for permitted use; and be accessible through repeatable and secure processing workflows.

AI readiness should also support reproducible training, testing and validation across institutions and populations, so that successful models can move beyond single-site pilots. Otherwise, Europe risks building powerful AI infrastructure without sufficient access to usable health data, while health data spaces remain well governed but underused.

Please note that there is no single architecture for connecting health data spaces and AI compute. Depending on the use case, either (i) authorised data can be moved to a secure compute environment, (ii) approved AI/analytics workloads can be brought to the data and executed locally or federated, or (iii) data can reside in shared secure cloud enclaves with compute co-located. In this third case in particular, the storage/compute infrastructure itself becomes strategic public-interest infrastructure. Such infrastructure should be under European control, potentially publicly funded or operated in public-private partnerships, and designed around open interfaces, portability and provider substitutability to avoid single vendor dependency and vendor lock-in.

A sustainable model therefore requires a shift from project logic to infrastructure logic. Europe and Member States should treat core health-data-space capabilities such as metadata services, onboarding, trust services, secure processing, terminology support and interoperability testing as reusable public-interest infrastructure. This reduces duplication, lowers entry barriers for hospitals and SMEs, and creates a basis on which both public-interest and commercial use cases can scale.

Call to action

The next phase for health data spaces cannot consist of another fragmented round of pilots. Europe now needs a coordinated effort to establish trusted, interoperable and sustainable health data spaces across Member States, while preserving national responsibilities and institutional autonomy. The paper translates this ambition into four mutually reinforcing recommendations.

First, Europe must **scale trusted health data spaces across borders**. Existing initiatives need to converge around shared objectives and connect their infrastructures, services and data resources. Making data jointly usable across Member States will create the scale and diversity required for healthcare, research and innovation. The ecosystem must also support controlled and efficient data sharing with like-minded countries beyond the European Union.

Second, Europe must **establish a minimum common interoperability and trust baseline**. Interoperability must cover the legal, organisational, semantic, technical and operational dimensions of data sharing. Participants need reliable digital identities, reusable trust mechanisms and standard contractual clauses that are recognised across Member States. This common baseline must also facilitate exchanges with other sectoral data spaces, e.g. those covering socio-economic and environmental data.

Third, Europe must **create durable funding and operating models**. Health data spaces need to move beyond temporary project funding and become sustainable infrastructures with viable long-

term operating models. This requires clarity on how common services, data curation, onboarding, compliance, trust mechanisms and secure processing are financed and maintained. These models must also safeguard European digital sovereignty and reduce critical dependencies.

Fourth, Europe must **connect health data spaces to its AI, innovation and research agenda**. Data-sharing mechanisms need to meet the requirements of the AI era and connect with AI Factories, data labs, European Digital Infrastructure Consortia (EDIC) and other European research and computing infrastructures. Trusted access to interoperable, high-quality and representative health data will provide the foundation for vertical AI applications in healthcare, research and public health.

Success will be measured by the extent to which health data spaces become an integral part of health data sharing and health-AI development across Europe and beyond. This includes increasing the number and diversity of datasets reused, connecting key European infrastructures through trustworthy data-sharing services, and involving all EU Member States in cross-border exchanges. Success also means implementing health-data use cases more rapidly and at lower cost.

Through these recommendations, the authors commit to bringing currently fragmented initiatives together in more coordinated actions that create scale across institutions, sectors and national borders.

Summary of the four recommendations:

Recommendation	Core objective	Principal actions
1. Scale trusted health data spaces across borders	Achieve sufficient European and international scale	Converge existing initiatives; federating access to data across Member States; enable controlled sharing with like-minded countries
2. Establish a common interoperability and trust baseline	Enable trusted exchanges between independently governed ecosystems	Ensure legal, organisational, semantic, technical and operational interoperability; use digital identities and recognised contractual clauses; support cross-domain exchanges
3. Create durable funding and operating models	Turn time-limited projects into sustainable infrastructure	Establish viable financing and operating models; maintain common services; safeguard European digital sovereignty
4. Connect health data spaces to AI, innovation and research	Make trusted health data available for the AI era	Adapt data exchange to AI requirements; connect with AI Factories and data labs; enable vertical health-AI applications

In conclusion, starting from the EHDS as a given, this paper presents several concrete and thought-through recommendations to make the scaling and AI ambitions realistic. The authors pledge to converge the ongoing fragmented initiatives towards more concerted actions, with an ability to create scale, within and across