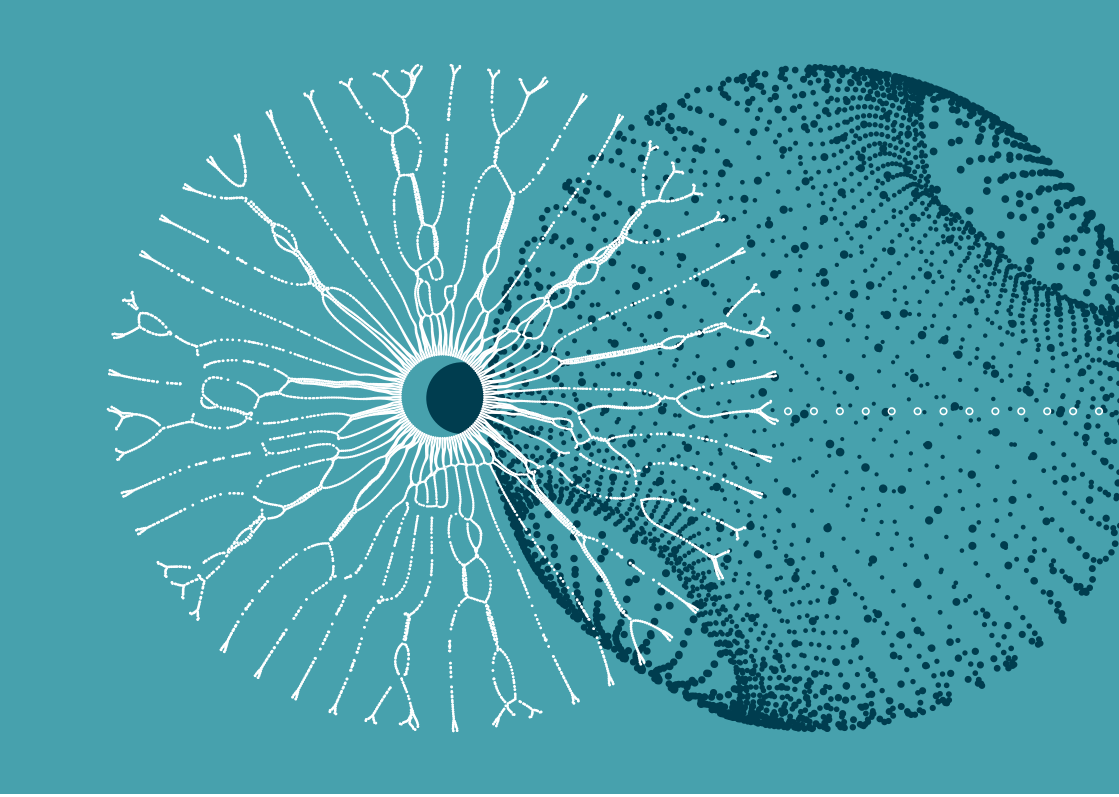


Embracing EHDS

How evolving your data strategy can help you prepare for the interconnected future of healthcare

/thoughtworks

Strategy. Design. Engineering.

A new era for health data sharing has begun

Over the past two decades, healthcare systems and organizations across the European Union have undergone a major digital evolution — harnessing the power of data to transform everything from research and drug discovery to diagnostics and care delivery.

However, every organization has transformed in its own way, at its own pace, creating a fragmented healthcare data environment and making it difficult to share data between organizations and countries. As a result, many organizations are burdened by inefficiencies introduced by data silos, and are unable to unlock the full value of all their data.

Now, that's all set to change thanks to the creation of the European Health Data Space (EHDS), which aims to establish a common legal framework and central cross border infrastructure (such as HealthData@EU pilot project) for the use of health data across the EU.

The EHDS originates from the European strategy for data laid out by the European Commission, which proposed the establishment of domain-specific common European data spaces to help solve challenges faced across the EU. The EHDS is the first domain-specific proposal designed to address health-specific challenges, especially those around electronic health data access and sharing.

Devised during the pandemic — which highlighted society's need for rapid healthcare data and insight sharing — the EHDS has been created to streamline research, innovation, policymaking, and regulatory activities across the EU, while giving patients greater control over their personal health data. The framework also provides new impetus for organizations to address some of their most persistent data challenges, such as breaking down data silos, and a strong business case for doing so.

If organizations and member states implement and adopt EHDS in the right way, it has the potential to ensure continuity of care, access to safe and high-quality healthcare, massively accelerate innovation, transform experiences for patients and, and lead to some of the biggest operational efficiency gains most healthcare organizations will have ever seen. However, to do that successfully, they need the right data strategy.

In this whitepaper, we'll explore what EHDS can help healthcare and life sciences organizations achieve, share practice evidence-based advice for overcoming the barriers to effective adoption, and conclude with an actionable checklist that will help your organization make the most of this once-in-a-generation opportunity.

Your EHDS opportunity

The EHDS will unlock the full value of primary and secondary health data use cases. Primary use cases relate to patients' digital control of their healthcare data and the portability of that data, and secondary ones relate to the use of data for research, innovation, and legislative policy creation.

Putting patients in control of their healthcare data

The EHDS gives patients by default the right and ability to share their health data themselves — deciding where and how it is used. That means it's about to become far easier for patients to undertake care journeys that span multiple organizations and countries.

For example, if a patient previously needed to receive care while traveling to another EU country, they may have needed to wait days just for relevant data about their care to be shared, oftenin paper-based formats. Now, information can immediately be passed on by the patient, enabling care delivery to begin right away.

An active choice will be given to the patients by the opt-out mechanism that allows natural persons to explicitly express their wish to not have their data processed for some or all secondary use purposes.

That isn't just great news for patients. Delays in sourcing and sharing patient data between organizations lead to large administrative burdens, missed appointments and lapsed care journeys. With greater access to data — driven and controlled by patients themselves — healthcare organizations will see a huge increase in efficiency and the quality of care they can deliver at speed.

EHDS facilitates the sharing and exchange of various types of data, going beyond just Electronic Health Records (EHRs). These include both personal and non-personal electronic health data, which encompasses a wide array of information related to health and well-being.

The first stream of priority categories for the primary exchange includes:

- **Patient summaries** in a harmonized format, including a minimal shareable set that can be expanded to disease-specific data. The patient summary will contain personal details, contact information, insurance, allergies, medical alerts, vaccination/prophylaxis information, medical history, medical devices and implants, procedures, but also patient-provided data, observation results, plan of care, and information on rare diseases.
- **Electronic prescriptions**
- **Electronic dispensations**

The second group of priority categories will follow up from 2029, scheduled to contain:

- **Medical images and image reports**
- **Laboratory results** obtained via in vitro diagnostics, including reports, tests and diagnostic results
- **Patient discharge reports**

This will be followed by the selected categories for secondary use related to genomics, omics and wellness applications.

Together, the categories of data will offer a general coverage of most health related data:

Genomic Data	Data related to the inherited or acquired genetic characteristics of a natural person, providing unique information about their physiology or health, often resulting from biological sample analysis. This can contain human genetic, genomic, and proteomic data.
Data impacting on health	This includes social, environmental, and behavioral determinants of health. For example, data on consumption of different substances, homelessness, health insurance, minimum income, professional status, behavior, and environmental factors like pollution and radiation.
Health-related Administrative Data	This encompasses claims and reimbursement data.
Person-Generated Electronic Health Data	Data from medical devices, wellness applications, and other digital health applications or wearables.
Pathogen Genomic Data	Includes pathogen genomic data impacting on human health. Access to and sharing of such data is essential for developing detection tools, medical countermeasures, and responses to public health threats.
Identification Data	Data related to health professionals involved in the treatment of a natural person.
Population-Wide Health Data Registries	Public health registries are included.

Data from Medical Registries	Electronic health data from medical registries for specific diseases.
Clinical Trial Data	Electronic health data derived from clinical trials.
Data from Medical Devices and Medicinal Products Registries	Electronic health data from medical devices and registries for medicinal products and medical devices.
Research Cohorts, Questionnaires, and Surveys	Data from research cohorts, questionnaires, and surveys related to health.
Biobank and Dedicated Database Data	Electronic health data originating from biobanks and dedicated databases.
Data Related to Status	This incorporates electronic health data pertaining to insurance status, professional status, education, lifestyle, wellness, and behavior data relevant to health.
Non-Personal Electronic Health Data	Data concerning health and genetic data in electronic format that falls outside the definition of personal data.

Data users can add various improvements such as corrections, annotations, enrichments to existing data sets following processing based on a data permit, resulting in enrichment of existing data sets.

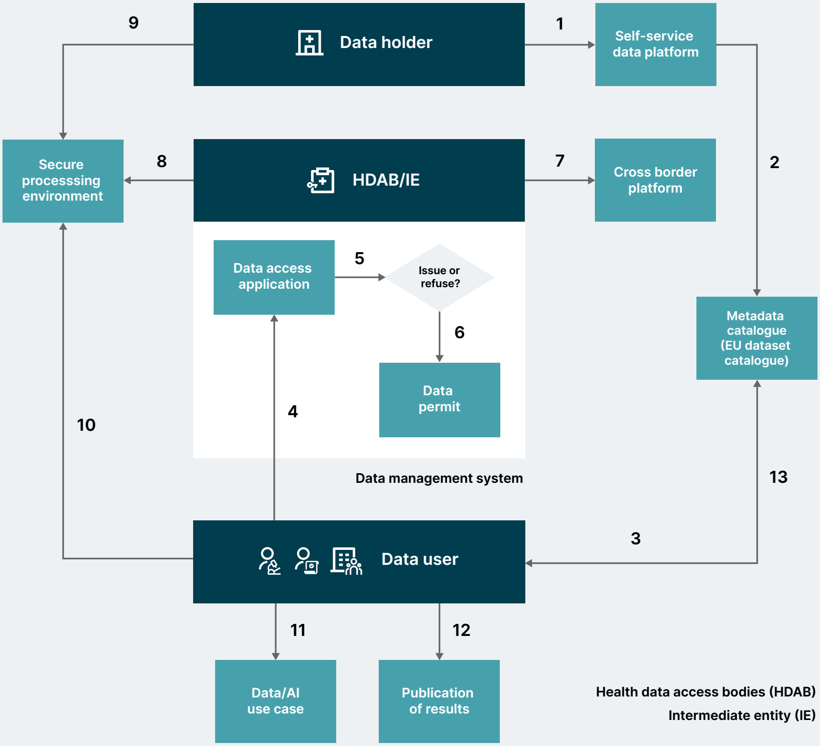
Transforming research, development, policy making, and healthcare and pharmaceutical innovation

For the first time ever, EHDS will give organizations easy, regulated access to data sets within and across borders for commercial use. For healthcare and life sciences organizations, that equates to three significant benefits:

- A massive increase in the volume, timeliness and diversity of data available for research, development and innovation purposes.
- The ability to generate more valuable research insights faster, improve responsiveness to new market demands and accelerate time to market.
- Cross-border access to data, helping organizations understand patients internationally and tailor their offerings for local populations.

That data can also be harnessed by policymakers to help them set more suitable standards for healthcare delivery and develop a deeper understanding of population health.

How will data users, data holders and health data authority bodies interact in the context of data exchange for secondary usage and which are the various systems and mechanisms involved in the process?



1. Data holder registers data sets to be available for secondary use using a HDAB governed account within a provided UI.
2. A registration workflow is triggered (potentially with HDAB approval / intervention) for metadata and governance policies related to the dataset. The metadata will be replicated at national dataset catalogues as well as EU wide dataset catalogue.

3. A data user (eg: a researcher, policy maker) browses data sets using the metadata catalogue.
4. A data user submits a data access application request using a data management system.
5. HDAB processes the data access application, looking at purpose, ethical usage, data minimization, ability to use the SPE. The result will be an issuance or refusal of a data permit.
6. In case of an issuance, a data permit is created and both data holder / data user are notified.
7. Cross border data information about available datasets and cross-border data access applications happens using cross border platforms, such as MyHealthData@EU.
8. HDAB grants access to data using a SPE.
9. Data holder uploads data in the SPE.
10. The data user accesses anonymized data using the SPE.
11. Data users develop a use case based on the data sets acquired within this flow, but often with a combination of data obtained from other sources (internal, external).
12. Data users should acknowledge and publish the research results when EHDS data sets have been used.
13. Data users can perform data sets enrichments, through correction, annotation, merges, additions, linking to other data sets or labelling.

Note: in a single data holder scenario, steps 5, 6 and 8 may fall into the data holders responsibility, and not the HDAB.

For smaller data holders (such as GPs), that do not have the necessary resources to cover the responsibilities related to this workflow, there will be Intermediary entities to help with the data and data exchange related tasks.

The final use case or data product at the data user will combine multiple sources of data, from internal data, open source data sets to commercial data and data obtained in the above described workflow.

A data management system will be used to record all data access applications, requests, permits. Additionally, permits must be publicly available.

Despite a large volume of medical and health data related being available or in the process of being entered into the system, there are multiple concerns related to the data quality and propagating issues related to the quality into the workflow. Understanding where data quality can be compromised and where bias is introduced is crucial. Data quality degradation and bias can arise at multiple points in the data journey.

- If the data entering the self-service platform is of low quality, all subsequent data products and analyses will be unreliable.
- Incomplete or inaccurate metadata in the national and european data catalogs will severely hinder the discoverability and understanding of datasets.
- As the enrichment process will enhance datasets with data originating from other datasets, potential low quality data sets can affect existing data sets. Auditability and provenance of data via enrichment are also challenges.
- The data access application process can also introduce bottlenecks and biases, potentially delaying or preventing legitimate research applications.
- Data users working with poor quality data or misunderstanding the data’s context will inevitably produce flawed analyses and incorrect conclusions.

- Transferring low-quality data across borders will amplify problems and hinder cross-national insights.
- The use of biased or low-quality data in AI and the publication of results based on flawed data will lead to the dissemination of misinformation, damage credibility, and perpetuate health inequalities.
- Pre-existing biases in the data and in the data access application process are a fundamental risk to equitable outcomes.
- It's important to proactively inquire about how biases can be introduced and consider their effects across different groups.

When EHDS is adopted in the right ways, everyone wins. Healthcare organizations can deliver a higher standard of care. Researchers can uncover valuable insights at unprecedented speed and train powerful AI algorithms using large, data sets labelled for quality. Pharmaceutical companies can develop innovative drugs and therapies faster. And the entire healthcare ecosystem across the EU becomes better informed, data-driven and more responsive to shifting population needs.

What’s standing in organizations’ way?

The EHDS represents great opportunities for healthcare and pharmaceutical organizations. But it also brings a new wave of challenges they must tackle to make the most of those opportunities.

The EHDS is laying a new foundation for decades of data sharing across the EU. It will require changes to all systems which interact with Electronic Health Record systems, impacting everything from wellness applications and emerging medtech to core applications utilized in hospital management.

With many healthcare institutions burdened by complex legacy systems and aging infrastructure, and levels of digital maturity highly varied between organizations and EU nations, not everyone will be ready to adopt the EHDS right away.

Key challenge: Delivering interoperability across a diverse landscape

The EHDS does not mandate digitalization. It instead regulates data once it has been digitized, giving member states flexibility in how they achieve new standards. As such, different levels of digitalization and interoperability will still need to coexist across the EU.

Core interoperability principles such as openness, availability of interfaces, technological neutrality and data portability, together with clear and precise terminology and definitions will guide the usage and adoption of medical standards and interoperability usage. This will be driven by common models such as OMOP, as well as common languages such as HL7-FHIR, DICOM, Open-FHIR, SNOMED, LONIC, DCAT-AP and others which show a large variation in usage and adoption.

By providing a common format to health data, common data models such as OMOP CDM make it easier to query, analyze, and compare data across different databases and institutions. However, CDMs adoption encounters challenges related to mapping and transformation, the translation from locally used standards to shared vocabulary, and the need of specialised personnel. The mapping and transformation process could lead to data quality issues.

This presents a significant challenge for organizations in the space. Varied levels of digitalization and interoperability both within and across member states — ranging from solitary centralized national systems to multiple centralized or decentralized approaches — is a major barrier to the seamless sharing of data regardless of location.

For example:

- **In Germany,** organizations have adopted a patient-moderated approach to EHRs, where patients have sole control over data access and sharing. However, its decentralized approach and low adoption of EHRs presents significant data sharing challenges.
- **In France,** the Health Data Hub initiative has created a high degree of centralization, and created an environment where data sharing is relatively simple and accessible.
- **In Spain,** teams are implementing EHDS using a collaborative model. The central government sets the direction and allocates funding, while autonomous regions have flexibility in implementation, which might lead to diversity of final implementation between individual regions.

Before organizations can leap in, they’ll need to ensure three key things:

1

High data quality and utility awareness

EHDS may enable an organization to share and access more data than ever before. But for most healthcare organizations, not all of their data is truly worth sharing today.

The EHDS establishes national and EU-wide data catalogues for easy data set discoverability and access. As such, data sets require quality and utility labels, so organizations can select the data that best suits their needs. The data sets might also get connected to other initiatives, such as the official portal for European data (data.europa.eu) to reduce the redundancy of data sets.

All data requires the right supporting documentation, including meta-data, model information, data dictionaries, information on standards used, and data provenance. The quality of their data will also need to be proven, demonstrating its technical quality, and showing its completeness, uniqueness, accuracy, validity, timeliness, and the consistency of the data.

While the data quality and utility label may provide a multi dimensional approach to the data quality, by introducing additional aspects in areas such as data enrichments or coverage, it is only an incipient project that provides dimensions and ways to measure these various dimensions of health data quality, such as completeness. However, it can only be applied at the data set levels, after the collection, translation processes and does not address the root causes of health data quality, nor do they promote any proactive measures for data quality.

Health data is often multi-source and heterogeneous, coming from various systems and in different formats. Multi-source and heterogeneous data requires better integration and standardization.

Data can be inaccurate or incomplete due to factors in the collection and recording processes, thus necessitating improved data capture and validation processes.

It is frequently siloed within different organizations or departments, hindering a holistic view, and therefore calls for better data sharing and interoperability within healthcare.

There's often low interoperability, including semantic interoperability, making it difficult for systems to understand and exchange data.

A significant portion of health data is unstructured or semi-structured, requiring additional processing for analysis.

There can be limited access to external data sources, even when those sources could provide valuable context. Insufficient or lack of supporting documentation makes it harder to understand and use data correctly.

With challenges in detecting, monitoring, and maintaining data quality levels, as well as managing issues like latency, availability, and freshness, core data quality attributes must be enriched with additional dimensions of healthcare data.

For example, a multidisciplinary perspective can be an indication of the breadth of the data sources captured, while population representativity may show the demographic balance and can help with bias detection, ensuring - for example - that there is proper inclusion on research studies. If a research dataset predominantly includes one demographic group, findings might not be generalizable to the wider population, potentially exacerbating health disparities.

Additionally, metrics related to access and provisioning, will indicate the freshness of the data and how quickly newly collected information is incorporated into the available datasets, as well as the efficiency of the process from the moment a data permit is granted to when the data is actually made accessible to the data user.

Data quality
Accuracy, completeness, consistency, timeliness, validity, uniqueness

Data quality management
Maturity level, review and audit processes, biases examination

+

Data documentation
Metadata, support, documentation, data model, data dictionary, standards used, provenance

Coverage
Representativeness, geographic/spatial coverage, granularity, multi-disciplinary coverage

Data enrichments
Modifications using merge, update, connection to other data sets

Access and provision
Access approval to provision time, collection to dataset time

View on dimensions for data quality and utility

In practice, there's a huge amount of information relating to their data and data sets which organizations must be prepared to provide, including (but not limited to):

- Average timeframes in which a natural person appears in a dataset
- The time between the collection of electronic health data and its addition to a dataset
- Information on data enrichments, such as when data was merged and adding data to an existing dataset, including links with other datasets

The quality of the data at source (data holders) has an immediate connection to the quality of the research results at the data user. Poor data quality brings the risk of unfair AI outcomes, especially perpetuating existing inequalities at a larger scale.

For example, an acute kidney injury detection algorithm may underperform for women because of sex imbalance in its AI training data. The imbalance of the training data can be correlated to actual lower incidence in acute kidney injury in women, but can also be rooted in professional bias. Healthcare professionals may hold biases related to age, sexual identity, disability, and weight, all of which can affect the quality of care and the data generated. For example, older adults may be assumed to be unwilling to receive treatment, leading to different care decisions.

To deliver and maintain high quality data that's contextualized with all of the additional quality information, organizations will need mature data quality management processes — including data reviews, audits and detailed bias examination.

2

Interoperability between systems and data sets

The EHDS mandates that the technical design of EHR systems and related services should enable patients to make their health data available to a health professional of their choice. The interoperability of these systems should allow for selective sharing without technical barriers.

Unfortunately, decades of gradual digital transformation have left most healthcare organizations with patchwork systems with many data silos and heterogeneous standards of interoperability. Even after proposed standards, such as the European Electronic Health Record Exchange format (EHRxF), have been unveiled, adoption has been limited, resulting in a fragmented landscape not only within the EU, but even within individual nation states. If organizations can't freely access, move and migrate data between their own systems, it will be extremely challenging to port that data to other EHDS participants.

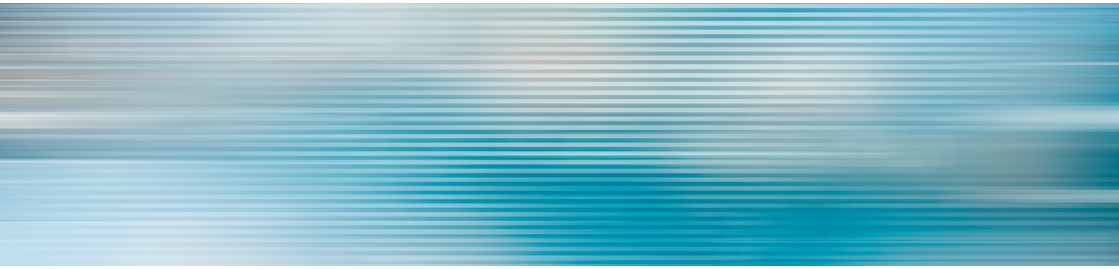
However, with the establishment of the EU Datasets Catalogue connecting the national catalogues of datasets established by the health data access bodies and other authorized participants in HealthData@EU, as well as the implementation acts, the member states must ensure all priority categories of personal electronic health data provided directly by a natural person or transmitted to a healthcare provider are using the specified format.

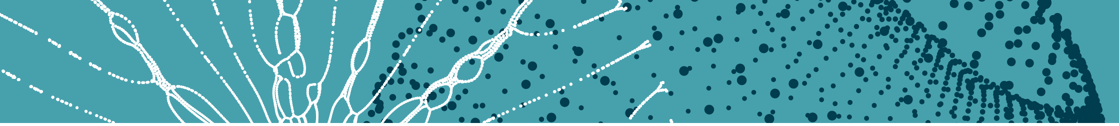
An important immediately impacted category are the **EHR manufacturers**, as EHDS introduce new obligations, including the mandatory self-certification scheme for EHR systems intended for the primary use of specific categories of electronic health data. EHR systems will need to comply with interoperability and security requirements, including personal data sharing and authentication mechanisms. Even “In-House” EHR systems must comply, taking on manufacturer obligations. This involves costs but also opens opportunities for accessing a larger, harmonized European market and fostering innovation. Costs include achieving interoperability, enhancing security, and undergoing certification.

3

A robust, scalable and secure infrastructure foundation

Healthcare data carries high governance and security demands — especially for personally identifiable data. So, in addition to enabling high interoperability and being scalable enough to meet future demands, organizations must ensure they can securely share it without exposing any personally identifiable information to anybody that should not have access to it. This is achieved through a combination of data minimization, anonymization, and pseudonymization. While international standards such as FHIR/HL7 are not explicitly part of EHDS, there is enough flexibility to reuse them, given they meet the general requirements for interoperability, security and data exchange formats.



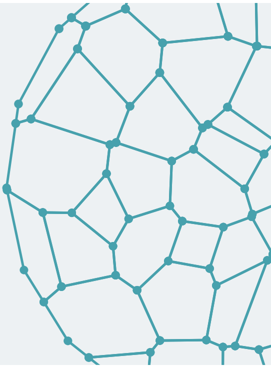


The solution — data strategy evolution

Unfortunately, most organizations today can't immediately embrace EHDS and gain all of its potential benefits, as they are hindered by a combination of all of the challenges mentioned previously.

What goes into a data strategy?

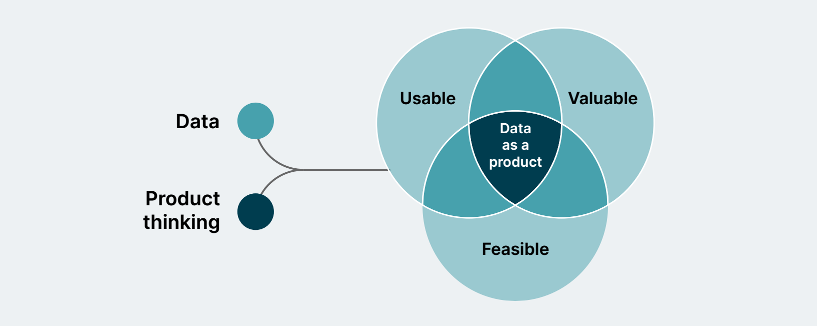
Your data strategy covers how your organization structures, manages, operationalizes, governs, and secures its data. A shift in your data strategy requires you to fundamentally evolve all of those areas — although not necessarily simultaneously.



To prepare for EDHS, organizations must first look inward to critically assess how they store, share and use data internally — and evolve their data strategy.

There is no single data strategy approach that will solve every challenge for every organization. However, the aims of the EHDS align very closely with the principles of data mesh and data as a product.

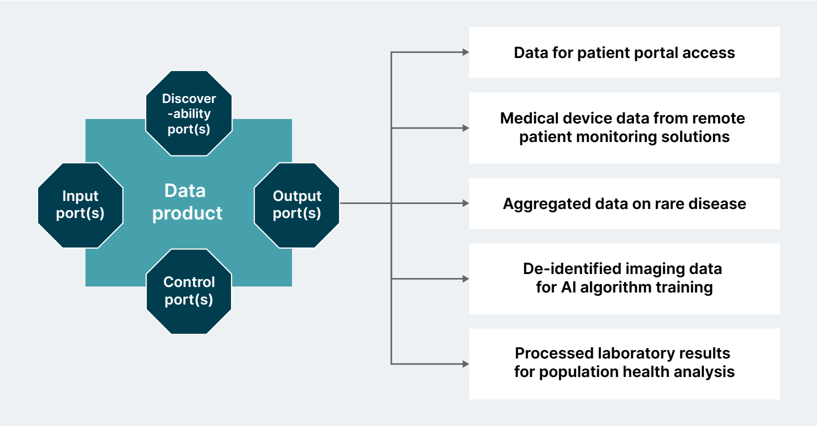
Data as a product applies product design and development principles to build high-quality, reusable and portable data solutions. Organizations often adopt it alongside self-service data platforms which remove friction and technological complexity from the interaction between data producers and consumers, and federated computational governance which automates data governance policies, with no need for a centralized authority.



Together, these principles and elements create environments where:

- Domains own, share and continuously improve the quality of data products built for their own use cases
- Governance is upheld by default, without limiting any team's ability to use data or share it with other domains and organizations
- People are empowered to find, access and innovate with the data they need, whenever they need it

In the context of the EHDS, they can help organizations build environments where data holders, users, and governing and access bodies can all easily securely share, access and control the same data sets.



As a self-contained unit of architecture combining code, data, metadata and policies, a data product also offers input and output ports, which aligns with the federated governance concept. Each of the output ports can be associated with specific governance policies and access controls determined by the data holder, EHDS regulations and Health Data Access Bodies (HDABs) based on the specific data access permit use case.

Additionally, the output ports of a data product can be designed to adhere to specific standards and exchange formats. Data products can be used both in the data holder organisation to prepare data for exposure, as well as the cross border federated exchange platforms to ensure proper sharing in the secure environment and in the satellite applications such as patient portals.

Building a foundation for rapid data sharing across the NHS

The NHS GP2GP record transfer service handles around 200,000 transfers between GP practices every month. In its efforts to continuously improve performance, find new ways to digitally support GPs and improve patient outcomes, NHS England identified this transfer process as an area of opportunity.

To help visualize the entire patient record management and transfer process in granular detail, and identify where the strongest opportunities for optimization and acceleration were, NHS England engaged Thoughtworks.

Following a series of collaborative discovery exercises, we began developing a dashboard using AWS-based infrastructure to help GP practices visualize GP2GP process performance. The dashboard enables GP practices and Clinical Commissioning Groups to clearly see how many transfers have taken place in their area, quickly spot any issues with ongoing transfers, and prioritize cases that require their attention.

To date, we've seen an 84% reduction in technical issues in the GP2GP system, which means far fewer records have had to be printed and transferred manually. The project has delivered an estimated year on year saving of 430,000 staff hours, equating to around £7.2 million in cost savings.

Crucially, this has also led to further process improvements and transformations across the NHS data landscape. We're now driving similar results in other processes and solution areas, helping to improve experiences for patients throughout England, and empower the frontline and back office teams that deliver and support care.

Now, let's take a closer look at some of the technical and cultural changes needed to bring that kind of data strategy and environment to life.

Establishing your foundation for evolution

The EHDS is the start of a new era in the use of healthcare data. So, it's important that organizations look beyond what they need today and build a scalable foundation that will serve them well for years to come.

For many organizations, this will be a powerful opportunity to critically assess their balance of on-premises and cloud infrastructure and migrate some of their siloed systems to a single cloud foundation. As laid out in the TEHDAS recommendations, cloud usage should be considered for the high scalability it offers, which can also help you optimize costs and accelerate time to market.

This decision shouldn't be made lightly. Your choice of cloud platform will ultimately impact:

- The scalability and performance of your data architecture, affecting the value that can be unlocked from your data use cases
- Your ability to access critical capabilities, such as those used to build self-service data platforms and other critical elements of your data environment
- The security of your data and ability to uphold robust governance and privacy standards as data is shared with other organizations

When we worked with one of the world's largest healthcare providers, Roche, to create a new data mesh foundation, we opted to build it on the Amazon Web Services (AWS) cloud. Throughout Roche's journey, the capabilities and scalability offered by AWS were instrumental in helping bring its vision for a highly connected, efficient and future-ready data mesh to life.

You can read more about Roche's journey in our Data Mesh in Practice article series, which explains the infrastructure and capability choices made within AWS and the impact they had on diverse users across Roche's domains.

Improving interoperability, portability and data quality with data as a product

Data as a Product (DaaP) is one of the core pillars of data mesh. It involves viewing data not just as a static resource, but as a valuable product that meets specific needs, solves explicit problems and provides tangible value to its consumers.

Just like a physical product, a data product is valuable and fully usable on its own, and encompasses all of the components essential for its use, such as metadata, transformation code, governance policies and even infrastructure.

When data is treated as a product, it:

- Is converted into valuable information, so it provides clear direction for problem-solving, decision-making and action
- Is presented to consumers in a secure, consistent, understandable and readily accessible manner, so their needs are addressed quickly and efficiently
- Meets consumers’ expectations regarding accuracy, completeness and timeliness, so that they can use the data with confidence
- Is Discoverable, Addressable, Trustworthy, Self-Describing, Interoperable, and Secure (DATSIS), so that it can quickly be integrated across the enterprise and is scalable with evolving business requirements

DaaP goes hand in hand with the concept of federated ownership and data governance. Data products are owned by the domains closest to them. They’re built to meet specific needs for that domain, but remain discoverable and immediately accessible for any other domain that wishes to utilize them for their own data use cases.

In the context of the EHDS, DaaP helps organizations:

- Transform how data is owned and managed across domains and put the people closest to data sets in control of them, improving the quality and relevance of that data
- Empower the data originators to take proactive measurements for data quality
- Build bespoke use-case based data products such as “patient-related data products” containing anonymized patient records, accessible to authorized researchers for clinical trials or population health studies, or “genomic data products” with standardized interfaces for sharing genetic information securely with other institutions
- Embed a culture of continuous data quality improvement, leading to better outcomes from an organization’s data use cases
- Package data with all the components required for its secure consumption, making it highly portable and shareable both internally and externally

The power of data products in practice

Thoughtworks has helped numerous organizations embrace the principle of Data as a Product, leading to significant operational and patient impacts:

At the United States Department of Veteran Affairs we created a data product designed to transform how the team calculated and monitors patient Medication Possession Ratios (MPR). The MPR calculation provides clinicians a view into how well a patient is adhering to a medication regimen — critical for veterans taking controlled medications like opioids.

The MPR data product provided a verified, trusted single source of truth for calculating MPR across VA. The processing time decreased by 88% and the iteration time to deploy new features was reduced by moving from a manual process to an automated, Continuous Delivery based process, allowing for new features to be delivered quickly and easily. And most importantly, the MPR data product acted as a template for further new data products designed to solve other patient and operational challenges.

Maximizing availability with self-service data platforms

In the EHDS, data permits give named data users permission to process electronic health data for specific secondary use purposes under certain conditions, served to them through secure data environments. However, at any time, any natural or legal person can submit a request for pre-processed anonymized statistical data.

Self-service data platforms can help organizations meet the needs of both data permit holders, and people who lodge data requests. Additionally, they give data holders within organizations on-demand access to all relevant data sets and products and empower them to quickly apply data to new use cases and operationalize it.

These platforms can support EHDS adoption by:

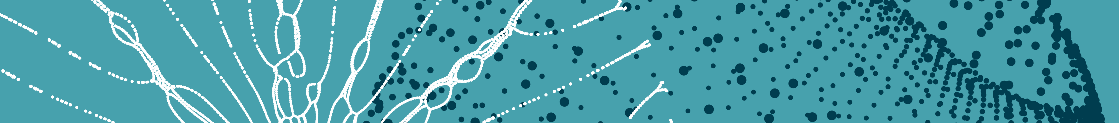
- Improving data sharing within organizations and breaking down silos that prevent data from being discovered, maintained and turned into value at scale
- Supporting interorganizational data sharing and making it easier to manage who can access data products and how they can be used
- Defining access, use and governance standards at the platform level to ensure all data products on the platform remain portable and interoperable
- Increasing the visibility, accessibility and discoverability of data, so that when someone needs something, they can get it quickly
- Removing friction and complexity from the interactions between data producers and consumers, enabling simple EHDS data sharing experiences

Connecting insights to action with a self-service data platform at Roche

Within the Data Mesh, the data platform has multiple planes. One common mistake that many organizations make is only focusing on the data infrastructure plane when devising and constructing a platform.

But for a Data Mesh implementation to be successful, teams need to carefully assess and make the right decisions at two further levels: the data product developer experience level and the mesh supervision level.

By carefully considering all three planes, and the capabilities required across them all during our data mesh implementation project at Roche, we were able to deliver a robust self-service data platform that seamlessly bridged the gap between data insights and human actions across diverse domains.



Federating governance and keeping data secure, but sharable

When domain teams create and own data products, they also take on responsibility for the governance and upkeep of those products. We call this federated data governance.

As a concept, it can be quite intimidating for organizations with robust governance demands. However, in practice, numerous governance guardrails implemented at the platform and product level ensure each domain adheres to the same standards.

This federated governance approach can help organizations maximize the value that they get from and contribute to the EHDS by:

- Supporting the secure sharing of data between organizations without exposing it to external risks — for example, through selective, role-based anonymization, while preserving data locality and sovereignty
- Improving the interoperability of data sets and products by ensuring they all meet a consistent standard
- Carefully balancing the high security and privacy demands of the healthcare sector with the need to freely access and innovate using data sets
- Restricting access to health data through a secure processing environment that complies with technical and security standards, often in combination with using various PET (Privacy Enhancing Technologies)

Our experiments with Anonymesh show how the privacy-utility trade-off can be addressed by employing a decentralized approach that keeps data at its source while facilitating secure information sharing. It also recognizes that there is no one-size-fits-all solution for data privacy, however the secure privacy environment must ensure that identifiable logs of access to verify and audit all processing operations are available and that security measures to mitigate potential security threats are in place.

Checklist: Preparing for EHDS and improving how you manage and utilize data

1

Start by looking inward at your own data challenges

To make the most of the opportunities the EHDS will create, your data, infrastructure, processes and governance frameworks all need to be of a high quality. The best way to determine where things may need to improve is by looking at the data challenges your teams are facing internally today. By finding your key points of friction and taking steps to resolve them, you'll make large strides toward preparing for the EHDS.

2

Critically assess your current infrastructure

If your infrastructure is comprised of diverse on-premises and cloud-based systems, with data siloed between them, it's worth taking this opportunity to break some of those silos down and establish a consistent, interoperable cloud foundation for the future of healthcare data sharing.

3

Prepare your people for change

Adopting a new data paradigm like DaaP isn't just a technical challenge. It's also a huge shift in culture and workflows for domain teams across your organization.

Make sure they know what's coming, how it will benefit them and what their new responsibilities will be as custodians of data products. Start with a clear vision of the data culture you want to enable across your organization, consult data holders and users on how they'd like that to be implemented, increase their data literacy and provide them with the skills they need to thrive in that new environment.

4

Focus on long-term ecosystem vision

Across the EHDS, multiple centralized and decentralized data architectures and approaches will need to coexist both locally and across the EU. Your focus should be on ensuring you can properly expose data as a data holder and connect to diverse systems today, and extend that connectivity as new systems as a data user emerge tomorrow.

5

Prioritize security and carefully evaluate risks

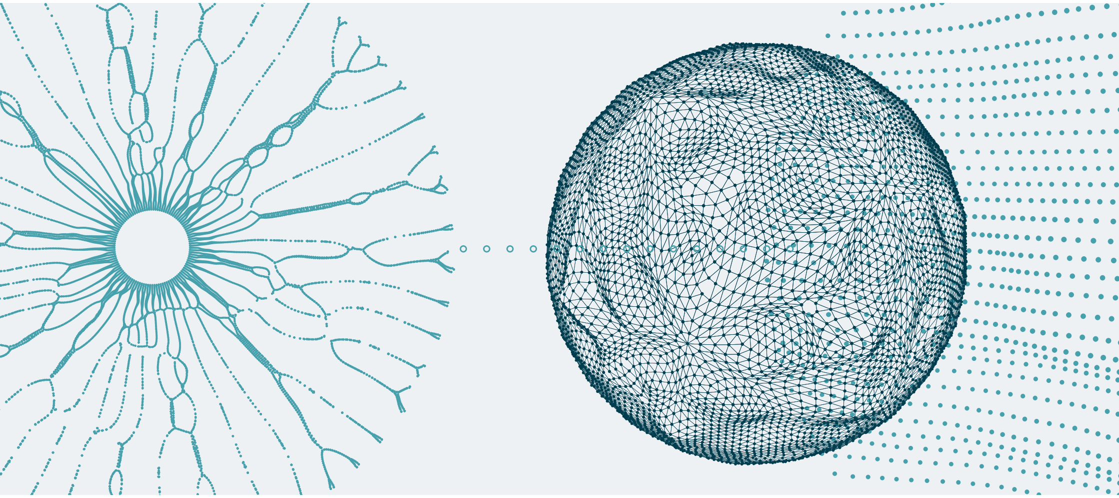
Data strategy evolution will bring a lot of change for your organization, so it's essential that all potential risks associated with that change are carefully modelled and evaluated.

Security should be a particular area of focus. Balancing the accessibility and sharability of data with the security demands of healthcare data is achievable through the application of DaaP, but requires careful management.

6 Think beyond the EHDS

For many healthcare and life sciences organizations, simply getting into a position to harness the full value-creation potential of the EHDS is ambitious. But while you're planning how to adopt and integrate into the EHDS, you also need to look far ahead and consider what you want the data-driven future of your organization to look like. This can involve taking advantage of diverse data sources, including medical synthetic generated data.

This is a unique opportunity to transform how you manage, structure, operationalize and generate value from your data. It's your chance to build the infrastructure and environment that will propel you to the forefront of your industry, and enable you to embrace powerful technologies like AI for many years to come.



The data-driven future of healthcare starts now

If your organization harnesses the full potential of the EHDS, you'll unlock opportunities to:

- Accelerate and enhance research and innovation processes, and bring valuable new drugs, therapies and products to market faster
- Transform outcomes for customers and patients and contribute to significant improvements in population health
- Streamline cross-border healthcare journeys and make it easier for everyone across the EU to access the care they need, where and when they need it
- Remove a massive amount of administrative work associated with the sharing of healthcare data and personal records
- Gain deeper insight than ever before into the most valuable healthcare data sets, and translate it into organizational, societal and patient value at speed

To do that, you'll need to evolve your data strategy and ensure you're ready for the new demands of the EHDS.

As the creators and leading implementers of data mesh, Thoughtworks has already helped numerous healthcare and pharmaceutical organizations navigate that journey. By applying data mesh's principles—such as DaaP—in unique ways for each organization, we help them create data strategies, infrastructure and environments that deliver real business value and impact for decades.

To learn more about the applications of DaaP and data mesh in the healthcare, pharmaceutical and life sciences industries, visit our website. Or talk to us today to explore how we could help your organization prepare for the EHDS and make the most of this unprecedented data opportunity.

Authors

Alexandra Lovin

Principal Technologist, Thoughtworks

In my over 15-year career as a software technology consultant, I've used my passion for technology to successfullyntranslate innovative technology strategies to reality, with proven ability to navigate complex challenges in diverse environments, from startups to enterprises. I am a resolute problem solver in both the tech arena and life at large, who finds solace in nature walks and creative writing.

Khrystyna Shlyakhtovska

Data and AI Program Manager for EHDS, AWS Data

In my 10+ year career as a program and business development professional, I've launched digital products for organizations in the financial sector and other industries. As a Data, AI and Industry Program Manager at AWS, I help customers - from startups to enterprises - navigate complex challenges related to data. One of my focus areas is coordinating projects that address customer needs in the emerging European Health Data Space. I am particularly excited about supporting our customers on their journey towards innovation in this evolving realm.

Thoughtworks is a global technology consultancy that integrates strategy, design and engineering to drive digital innovation. We are over 10,000 Thoughtworkers strong across 48 offices in 19 countries. For 30+ years, we've delivered extraordinary impact together with our clients by helping them solve complex business problems with technology as the differentiator.

thoughtworks.com

/thoughtworks

Strategy. Design. Engineering.