HEALTH DATA TALK SERIES:

Data Discovery: The Data User’s Journey according to the European Health Data Space regulation and HDA’s role in it

INTRODUCTION

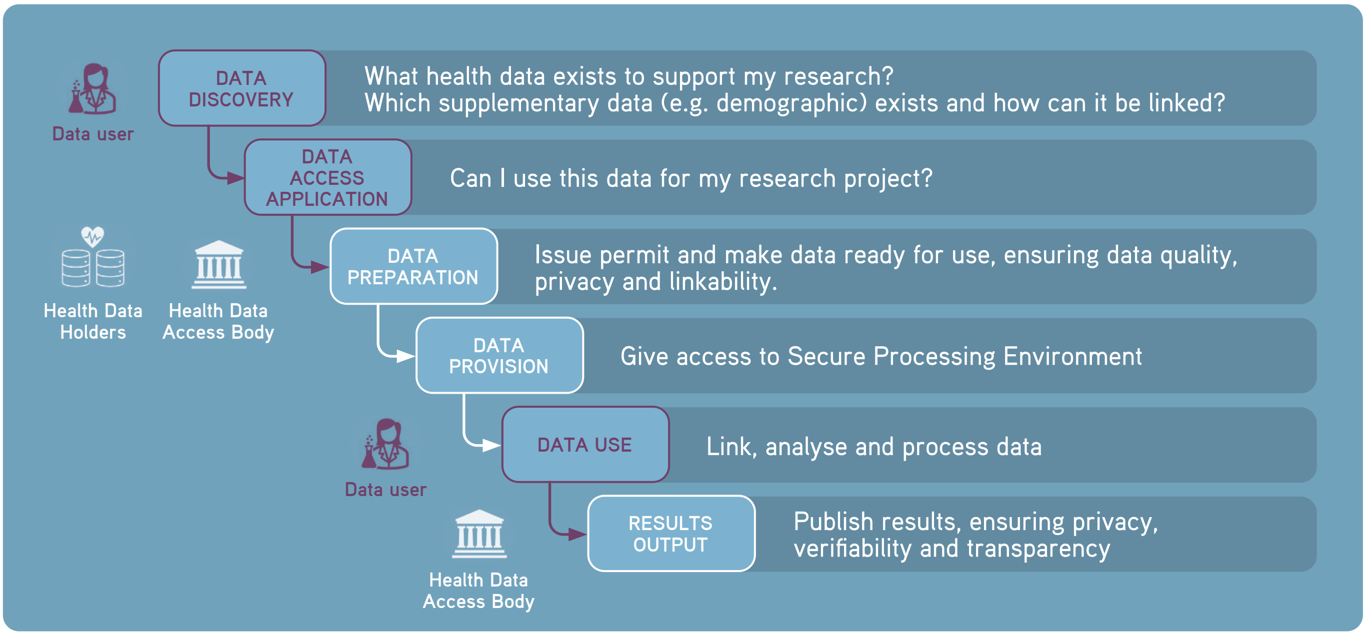
The European Health Data Space (EHDS) will introduce a major change in the way health and healthcare related data will be shared, processed and accessed for secondary use. This transformation will significantly affect the everyday work of both data holders (hospitals, registries, health care institutions…) as well as data users (researchers, policy-makers..). This article compares the health data user journey[1] before and after the EHDS regulation is in place, with an emphasis on the prescribed procedures, the obligatory privacy safeguards and the (future) role of the Belgian Health Data Agency (HDA) in this changing process.

USER JOURNEY BEFORE THE EHDS REGULATION[2]

Currently, a data user is obliged to search for the required data across different organizations (each acting as a “data holder”). Each organization needs to be contacted separately to find the data and to negotiate an agreement to receive the anonymized/pseudonymized data or an aggregated data analysis report. This data request process is very time- and effort- consuming, often requires substantial prior knowledge on the databases, and the final decision to release the data is dependent on the data holder and or whether the data is ready for re-use. Data exchange or transfer methods are - in some cases not secured and there is little to no control over the specific processing performed by the data user. Moreover, this might contribute to a high risk of re-identification of the patients whose health data is involved.

USER JOURNEY WITHIN THE EHDS

The implementation of the EHDS regulation, operational from 2029 onwards, will guide, facilitate and improve the data request process. The EHDS regulation imposes each Member State to install (at least one) Health Data Access Body (HDAB) that both facilitates and orchestrates the data request and access procedures, and issues a permit for data access if applicable.



What health data exists to support my research?
Which supplementary data (e.g. demographic) exists and how can it be linked?

Can I use this data for my research project?

Issue permit and make data ready for use, ensuring data quality, privacy and linkability.

Give access to Secure Processing Environment

Link, analyse and process data

Publish results, ensuring privacy, verifiability and transparency

Figure 1: The Data User Journey - Tehdas2 Joint Action

The user journey (see figure 1) starts from a **data discovery phase**. Instead of the need to contact different organizations separately to find the required data, each member state will provide a national contact point (e.g., the HDAB) that will set up a metadata catalogue with a description of all existing national health(care) related datasets.

---

[1] For a more elaborate description of the User Journey, please see the ‘Draft Guideline for data users on good application and access practice’ (THEDAS2), currently under review.

[2] For more information, see the Health Data Talk Series article “Data Linkages in Belgian (administrative) RWD: the missing link” by Degefaye Zelalem Anlay (Vrije Universiteit Brussel), on HDA Academy: https://academy.hda.belgium.be/course/view.php?id=55

Each national metadata catalogue will be connected to the European HealthData@EU platform, that will provide a metadata catalogue containing descriptions of all Member State metadata catalogues and European datasets.

HDA's role:

The Belgian national metadata catalogue developed by the HDA, describes existing health datasets in Belgium and is available via https://www.hda.belgium.be/en/data_catalog. Health-specific keywords and specific glossary functions are key features that ensure enhanced data discovery. The catalogue is also linking to other existing metadata catalogues, which helps data users to efficiently discover available data resources via one single point of contact. To provide a complete and user-friendly metadata catalogue, the HDA offers data holders technical, legal and semantic support in onboarding their datasets. Datasets are described in a standardized way using the Health DCAT-AP guidelines. Moreover, it includes variable-level information, providing data users immediately the required information to evaluate the usability of the data. The Belgian National Metadata Catalogue will be connected directly to the Healthdata@EU portal, enabling international data requests and the exchange of metadata to the EU catalogue.

After the discovery phase, the **data request process** follows to obtain a request approval or a permit application. Directly from the metadata catalogue, the data user can decide to submit a data request or a data access application to the HDAB. If the requested data are aggregated, statistical or non-individual level data, the request will be a "**data request form**" without the need for a permit. For access to anonymized or pseudonymized data at individual level, a "**data access application**" should be submitted to the HDAB that will perform a more extensive assessment of the application, including the needs and motivation for the use of anonymized or pseudonymized data at individual level. The HDAB will approve or refuse data access. In case of approval, a permit is granted specifying the conditions for data access (specific variables, access duration, minimization, pseudonymization, type of analysis).

HDA's role:

As the HDA is the **single point of contact** for data (access) requests, also across multiple data holders, the data user will only need to apply once for the required data. By evaluating data access applications, the HDA will ensure that **legal and ethical criteria** are met for access to sensitive data in accordance with EHDS rules. Moreover, the HDA will collaborate with the Information Security Committee[3] (ISC), to avoid the duplication of access applications. Together with the ISC, the HDA will handle data access applications as a one-stop shop in evaluating the application and delivering the permit. This will reduce fragmentation and duplication in the permit process, which will also shorten timelines for access.

During the **data processing stage**, the requested **data is prepared** and made available in accordance with the conditions specified in the permit. After these measures to respect and protect the citizens' privacy, the data is transferred and **made available in the** HDAB's **secure processing environment (SPE)**[4]**.** This SPE guarantees that only authorized persons have access to the data. The data user is allowed to access the data and perform the research within the SPE during the granted time period.

---

[3] For more information, see https://www.ksz-bcss.fgov.be/nl/page/ivc-in-het-kort

[4] For more information, see https://academy.hda.belgium.be/mod/glossary/showentry.php?eid=784

HDA's role:

By supporting the preparation and the provision of the data, the HDA will help data users by ensuring that permitted data is accessible in a **secure processing environment** and only by the authorised user, minimising re-identification risks. Based on the national social security identification number, the HDA will enable patient level linkage of data from different data sources. Moreover, the HDA will manage an automated filtering of individuals based on its opt-out system that ensures that Belgian citizens can easily opt-out from secondary use of their data.

After the data processing stage and the data user's research stage, the process ends with the **results reporting** phase. The **data user** is obliged to inform the HDAB about the results of the analysis and about any significant findings relevant for patients and critical errors in the data. The EHDS regulation imposes the **HDAB** to publish an overview of the requested datasets and its research outcomes to promote transparency and to build trust.

HDA's role :

The HDA will publish **a transparency portal**, providing an up-to-date registry of all approved and rejected data requests. This will ensure transparency about all data usage towards citizens. It thereby informs data users about previous data requests, avoiding duplication of requests.

By **publishing the research results** of the data users in an understandable language, the HDA will support both the healthcare community as well as citizens in raising awareness of the benefits and impact of the secondary use of data.

In summary, by centralising metadata and request management, the HDA drastically **reduces administrative burden** for data users, accelerates access timelines, **prevents duplication of effort** by making existing datasets discoverable and reusable, and **encourages standardisation** and shared documentation across the health data ecosystem. As part of the **HealthData@EU infrastructure**, the HDA also enables data discovery and access **across EU Member States**.

Signed by the Health Data Talk Series group

Hendrik De Rocker
(APB)

Inge Van de Velde
Stefan Gijssels
(Patient Expert Center)

Isabelle Huys
(KULeuven)

Katleen Janssens
(Health Data Agency)

Karen Crabbé
(pharma.be)

Nick Marly
(Cabinet public health)

Stefanie Devos
(beMedTech)

Sofie Bekaert
(King Baudouin Foundation)

Sofie De Broe
Wannes Van Hoof
(Sciensano)

Teodora Lalova-Spinks
(Ghent University, KULeuven)