

The Health Data Research Innovation Gateway: A platform to discover, request access to health data and demonstrate trustworthiness in research across the United Kingdom and beyond

Susheel Varma, Paola Quattroni, Chris Milner, Neil Postlethwaite, Ruth Milne, and Elaine Brannigan*

Health Data Research UK

**Ruth.Milne@hdruk.ac.uk*

ABSTRACT

With funding provided by UK Research and Innovation (1), Health Data Research UK – the national institute for health data science – (2) has built the Health Data Research Innovation Gateway (3) as an interoperable access point for metadata of health data resources in the United Kingdom. The Gateway is an integral part of HDR UK’s strategy (4) to bring together the UK’s health data custodians via the UK Health Data Research Alliance (5) and create specialist centres of excellence across the UK via the Health Data Research Hubs programme (6).

The “Gateway” was established in 2020 and since then over 750 datasets have been made available for researchers to search and request access to, in addition to over 3000 other health data resources including publications, data uses, tools and educational courses. The Gateway is not limited to being a metadata registry, but forms an entire ecosystem for fostering the curation, quality improvement, access, and ethical use of health data. To date the Gateway has over 2000 registered users from across the health data community who are making more than 10,000 searches a month.

Our agile approach to design and development of the Gateway, including continuous and frequent engagement with multiple stakeholders, is enabling us to build an infrastructure and service that’s open, federated, and interoperable with a clear user journey (Fig. 1). During the initial development phase (from mid-2020 to mid-2021) we were able to create a user-friendly portal with built-in reporting dashboards, make Gateway assets more discoverable via improvements to metadata onboarding, incorporate the Data Utility Framework (7) to support data curation, and harmonise the data access request management process for data custodians.

For IWSG 2022, we will present our development and operational progress from the last 12 months (from mid-2021 to mid-2022) which will focus on three areas of new Gateway functionality: Cohort Discovery (8); Trusted Research Environments (9); and the Data Use Register (10).

The Cohort Discovery search tool accelerates the journey to impactful research by allowing users to search by specific population criteria across multiple datasets at once. It uses the technology developed for ATLAS, an advanced data search tool for researchers (11); and builds on CO-CONNECT, a project led by the Universities of Nottingham and Dundee in collaboration with BC Platforms (12).

The integration of Trusted Research Environments – highly secure computing environments that provide remote access to health data for approved researchers – with the Gateway is necessary to deliver fast, safe, and responsible access to data that 1) accelerates the pace of quality research and 2) builds the confidence and trust of our stakeholders. This work is based on a paper published by the UK Health Data Research Alliance and the National Health Service Transformation Directorate (13).

The Data Use Register hosted by the Gateway shows how datasets published on Gateway are being used, by whom and for what purpose, supporting transparency in the use of health data across the community. The functionality implements a national standard for data use registers created by the UK Health Data Alliance (14) and aligns to the Five Safes framework (15). Data use reporting is also automatic for data custodians using the Gateway data access management system.

The Gateway is in constant development so our talk will consider our recent successes as well as our lessons learned to date. We will also share some insights and challenges about its future development, including the integration of international datasets (16, for example).

Keywords: *Innovation Gateway; Cohort Discovery; Trusted Research Environments; TREs; Data Use Registers; Health Data Research UK; HDR UK; UK Health Data Research Alliance*

REFERENCES

- [1] UK Research and Innovations 2022 < <https://www.ukri.org/> >
- [2] Health Data Research UK 2022 < <https://www.hdruk.ac.uk/> >
- [3] Health Data Research Innovation Gateway 2022 < <https://www.healthdatagateway.org/> >

- [4] HDR UK One Institute Strategy 2019 <<https://www.hdruk.ac.uk/wp-content/uploads/2019/11/191010-HDR-UK-One-Institute-Strategy-compressed-for-website.pdf>>
- [5] UK Health Data Research Alliance 2022 <<https://ukhealthdata.org/>>
- [6] Health Data Research Hubs, HDR UK 2022 <<https://www.hdruk.ac.uk/helping-with-health-data/health-data-research-hubs/>>
- [7] B. Gordon, J. Barrett, C. Fennessy, C. Cake, A. Milward, C. Irwin, M Jones, and N. Sebire “Development of a data utility framework to support effective health data curation” *BMJ Health & Care Informatics* 2021; doi: 10.1136/bmjhci-2020-100303
- [8] Health Data Research Innovation Gateway, Cohort Discovery 2022 <<https://www.healthdatagateway.org/about/cohort-discovery>>
- [9] HDR UK, Trusted Research Environments 2022 <<https://www.hdruk.ac.uk/access-to-health-data/trusted-research-environments/>>
- [10] HDR UK, Data Use Registers 2022 <<https://www.hdruk.ac.uk/access-to-health-data/data-use-registers/>>
- [11] ATLAS 2022 <<https://biobankinguk.org/advanced-search-tool-for-researchers-atlas/>>
- [12] CO-CONNECT <<https://co-connect.ac.uk/>>
- [13] S. Varma, T. Hubbard, D. Seymour, N. Brassington, and S. Madden “Building Trusted Research Environments: Principles and best practices towards TRE ecosystems” Zenodo 2021; doi: 10.5281/zenodo.5767586
- [14] N. Karrar, S. Khan, S. Manohar, P. Quattroni, D. Seymour, and S. Varma “Improving transparency in the use of health data for research: Recommendations for a data use register standard” Zenodo 2021; doi: 10.5281/zenodo.5084761
- [15] Five Safes, Office of National Statistics 2022 <<https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/>>
- [16] International COVID-19 Data Alliance 2022 <<https://icoda-research.org/>>

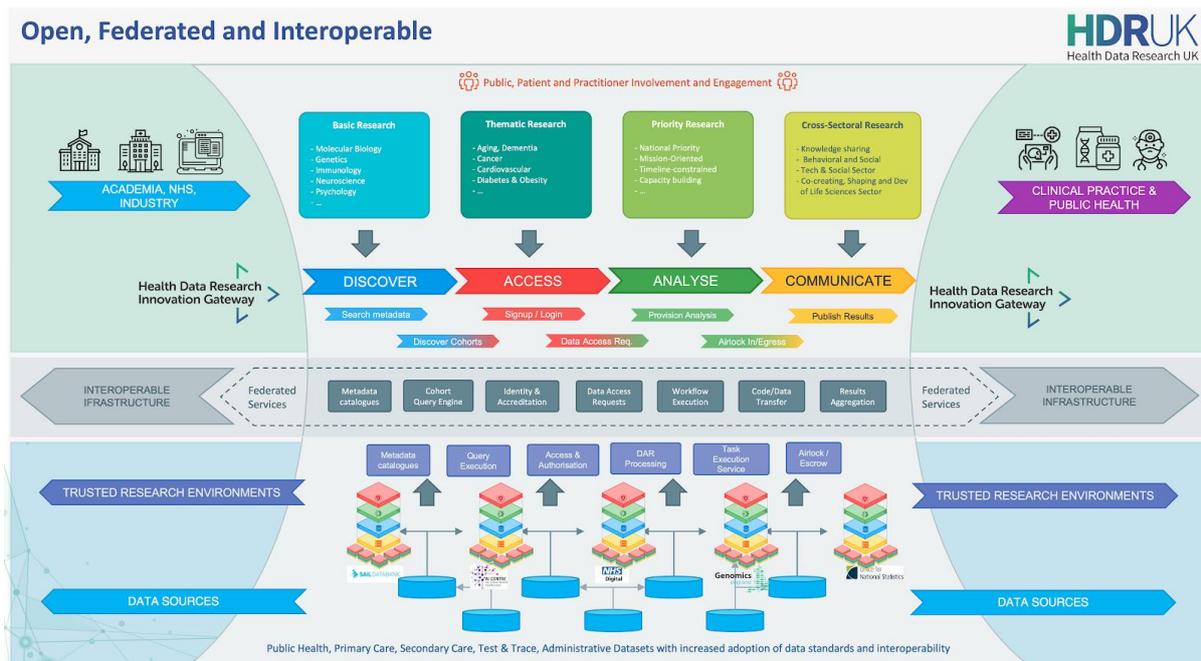


Fig. 1. The Health Data Research Innovation Gateway is developed using an infrastructure that’s open, federated, and interoperable with clearly defined phases in the user journey.