

**The Health Data Research Innovation  
Gateway: A platform to discover and request  
access to datasets, and demonstrate  
trustworthiness in health research across the  
UK and beyond...**

14<sup>th</sup> International Workshop on Science Gateways (IWSG)  
Trento, Italy  
15-17 June 2022



# Outline

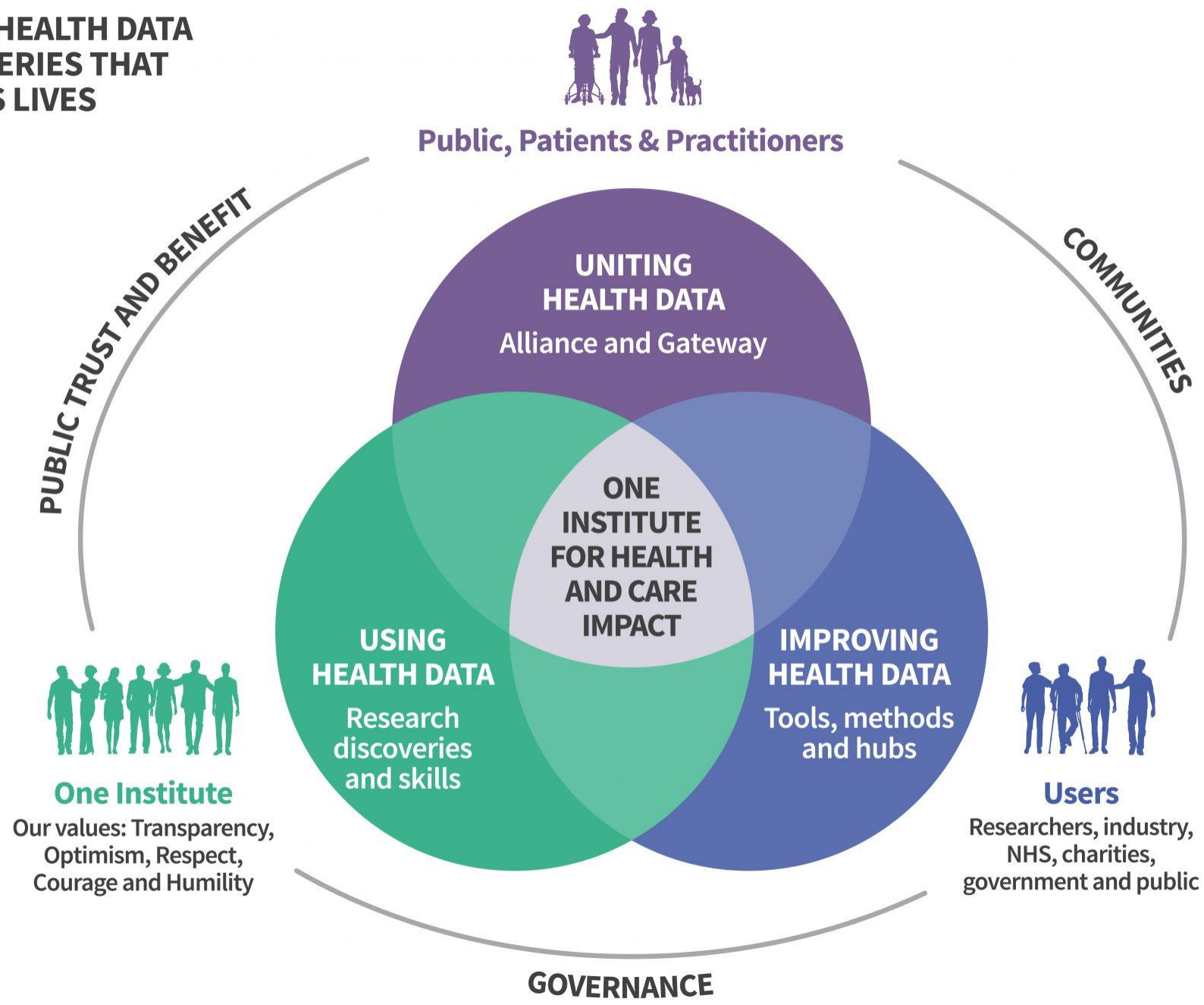
- Who we are and what we do
- An introduction to the Health Data Research Innovation Gateway
- Spotlight 1: Cohort Discovery
- Spotlight 2: the Data Use Register
- What have we learned and what's next?
- Summary and questions



**We are the national institute in the UK for health data science**

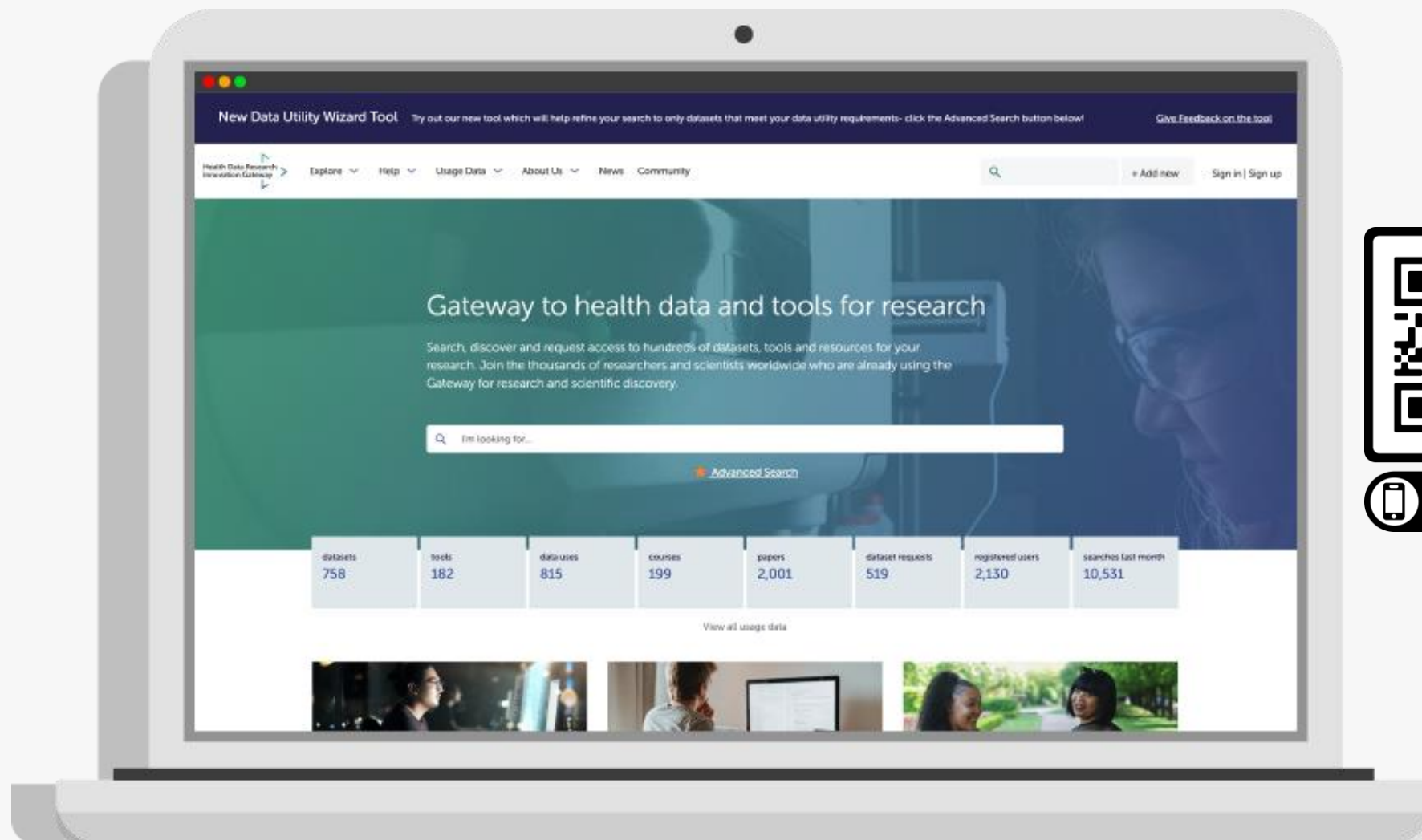
**Our mission is to unite the UK's health data to enable discoveries that improve people's lives.**

UNITING THE UK'S HEALTH DATA  
TO ENABLE DISCOVERIES THAT  
IMPROVE PEOPLE'S LIVES





[www.healthdatagateway.org](http://www.healthdatagateway.org)



# Health Data Research Innovation Gateway

The journey so far and the road ahead



Explore the timeline

**Launch of the Gateway Portal Minimum Viable Product** – the UK's first unified platform for discovery and access to health data

February

**HDR UK Phenotype Library** now integrated in the Gateway

September

**National Core Studies COVID-19 datasets** made available on the Gateway

November

**Trusted Research Environments (TREs)** collection now live on Gateway

December

2021

**First COVID-19 vaccine datasets** listed on the Gateway

March

**100,000** searches completed on the Gateway

**1,000** users registered

May

**600** datasets

**23,716** dataset views

**247** data requests

**April**  
**New Cohort Discovery Tool** enables users to search across multiple datasets

2022

January

**Data Use Register released**, an implementation of the UK Health Data Research Alliance data use register standard



**October**  
**Data Utility Wizard** released to improve data discovery



**140 datasets** now accessible via the **Gateway Five Safes Form** across eight data custodians

**230,100** searches completed on the Gateway

**2,100** users registered

**755** datasets

**150,815** dataset views

**511** dataset requests

**805** data uses

## Register for a Gateway account

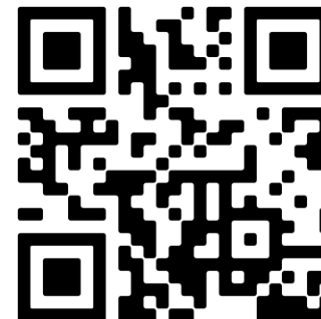
Search, discover and request access to hundreds of datasets available for research and scientific discovery

## Sign up to our monthly newsletter

Be the first to know about all the latest Gateway updates by subscribing to our monthly newsletter

Health Data Research Innovation Gateway > **HDRUK**  
Health Data Research UK

Connect with us



SCAN ME



**Federated metadata** – automatic import of datasets from other catalogues



**HDR Global** – integration of the ICODA Workbench to support the addition of global datasets to the Gateway

August



**Metadata version upgrade** – enriched dataset descriptions to improve search and discovery

December



**One data access request for multiple datasets** – to allow researchers to make a single data access application across multiple custodians

February

2023

September

**Data Use Register widget** – to allow custodians to display data from the Gateway Data Use Register on their website

July

**Streamlined Data Access Requests** for a self-service set-up for data custodians



**User interface updates** to improve the ability to select and filter search results

2022

June

# Cohort Discovery: meeting specific data needs

The Innovation Gateway's Cohort Discovery search tool allows researchers to search, assess and request access to potential datasets that precisely match information on particular populations of interest – for example data across multiple datasets relating to non-smoking patients aged 18 to 30 that have asthma and diabetes.

The tool enables multiple dataset searches to be done in a fast, secure, de-identified and ethical way. At no point is there any individual level data transferred from the custodian.



## How can Cohort Discovery help you?



Gateway users, including researchers, can carry out more specific searches and assessments on datasets listed in the Gateway, so that potential datasets which match the requirements of a research project are easier to find.



For data custodians, Cohort Discovery means requests to access data are more likely to result in success, whilst still retaining control over who has access to the data.



HDR UK has extensively tested the cybersecurity of the Cohort Discovery tool, and had this independently verified, providing confidence in its privacy and security.





# Data Use Registers

Getting the greatest use and benefit from the UK's rich health data is an important goal– but it is essential that those whose data is being used can see what is being done with it.

The Gateway data use register functionality clearly shows how datasets published on Gateway are being used, by whom and most importantly for what purpose.



## What benefits does the data use register offer?



**Automated process:** Data custodians who use the Gateway data access management system are automatically told about use of their datasets. This cuts down on administration time and provides the public with real-time data.



**Closing the loop on impact:** By providing a clear link between data use and research outputs we can better demonstrate the impact and value of using health data for research.



**Alignment to Five Safes Framework:** To show that only qualified people in secure settings can access the data, and for purposes that benefit the public, the content and structure of individual data uses follows the Fives Safes framework.



**Best practice:** The register implements a **national standard for data use registers**, developed after a wide consultation with data custodians and the public.

**View the Data Use Register**

For more information, visit the webpage

<https://www.hdruk.ac.uk/access-to-health-data/data-use-registers/>



## Our two biggest lessons



CHALLENGES IN MOVING FROM  
'START-UP' TO 'PRODUCTION'



BUILD THE *COMMUNITY* AND  
THEY WILL COME

# Health Data Research Innovation Gateway

The journey so far and the road ahead



Explore the timeline

**Launch of the Gateway Portal Minimum Viable Product** – the UK's first unified platform for discovery and access to health data

February

**HDR UK Phenotype Library** now integrated in the Gateway

September

**National Core Studies COVID-19 datasets** made available on the Gateway

November

**Trusted Research Environments (TREs)** collection now live on Gateway

December

**First COVID-19 vaccine datasets** listed on the Gateway

March

**100,000** searches completed on the Gateway

**1,000** users registered

May

**Data Use Register released**, an implementation of the UK Health Data Research Alliance data use register standard

January

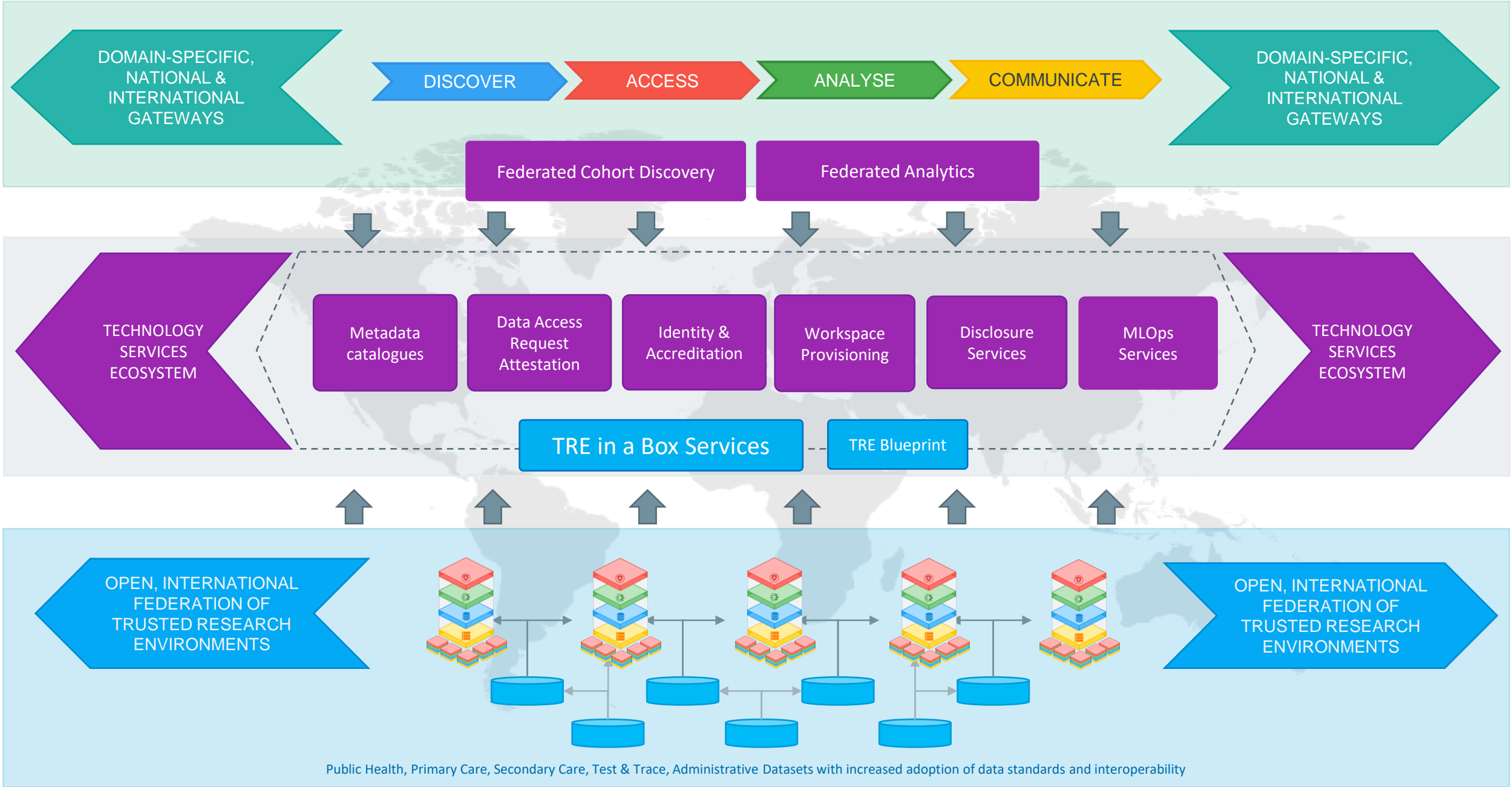
**230,100** searches completed on the Gateway

**2,100** users registered

**755** datasets

May







**Thank you!**

**Any questions?**



## Related resources



SUBSCRIBE TO OUR MONTHLY  
NEWSLETTER



CHECK OUT OUR PAST, PRESENT  
AND FUTURE EVENTS



LEARN MORE ABOUT THE HDR  
INNOVATION GATEWAY

Email: [ruth.milne@hdruk.ac.uk](mailto:ruth.milne@hdruk.ac.uk)  
Twitter: @HDR\_UK