Data Processing Service (DPS)

November 2020 Progress Update

Graham Spearing, Product Lead, Data Alliance, NHS Digital

Information and technology
for better health and care

Health Education England

Presentation Content

- NHS Digital – who are we?
- The Data Processing Service – progress and roadmap
- NHS Digital response to Covid-19

Our mission

To harness the power of information and technology to make health and care better

Our role

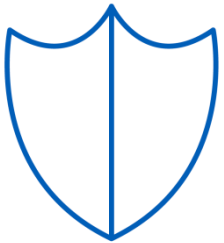
The national information and technology partner to the health and care system

NHS Digital's place in the health and care system

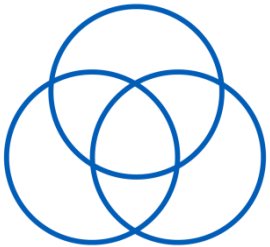
Our official name is the **Health and Social Care Information Centre** (HSCIC). We have used the name **NHS Digital** since July 2016. The Health and Social Care Act 2012 ordered our creation by statute, and sets out our responsibilities. The Department of Health (DH) set up the HSCIC in April 2013, as an executive non-departmental public body. We report to DH and implement DH policy. We work closely with NHS England, the National Information Board, NHS Improvement, and all parts of the NHS and social care in England.

- collecting, analysing and publishing health and care data
- making submitting data as easy as possible for health and care staff
- providing national technology for health and care services
- producing information standards
- improving the quality of health and care information and data
- publishing national indicators for health and care, to measure quality of care and progress against policy initiatives
- giving advice and support to health and care organisations on information and cyber security-19

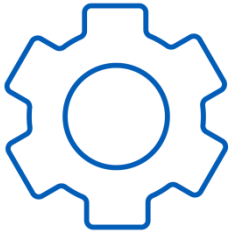
NHS Digital strategy



Ensure that every citizen's data is protected



Establish shared architecture and standards



Implement services that meet national and local needs

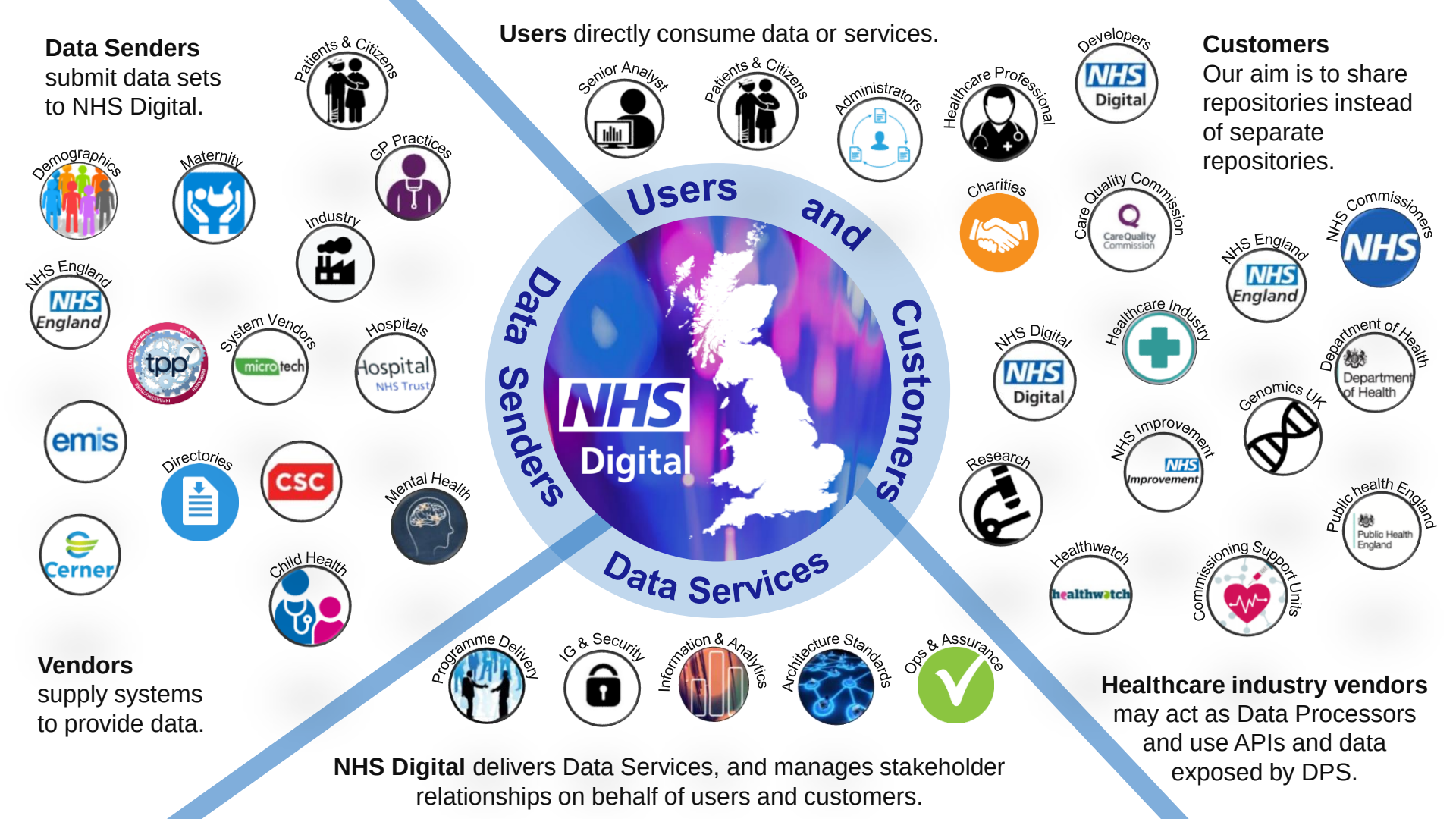
Help local organisations get the best from technology

Make better use of data and information

Data Senders
submit data sets
to NHS Digital.

Users directly consume data or services.

Customers
Our aim is to share repositories instead of separate repositories.



Vendors
supply systems
to provide data.

Healthcare industry vendors
may act as Data Processors and use APIs and data exposed by DPS.

NHS Digital delivers Data Services, and manages stakeholder relationships on behalf of users and customers.

Overview of our Data Processing Services

HEALTH AND CARE PROVIDERS

Data Services

Data Acquisition

Data Optimisation

Insights & Statistics

Data Access

Data Processing Services

Collect → **Process** → **Access**

Data capture and landing

Data processing, DQ, Bus. Rules, MPS and De-ID services

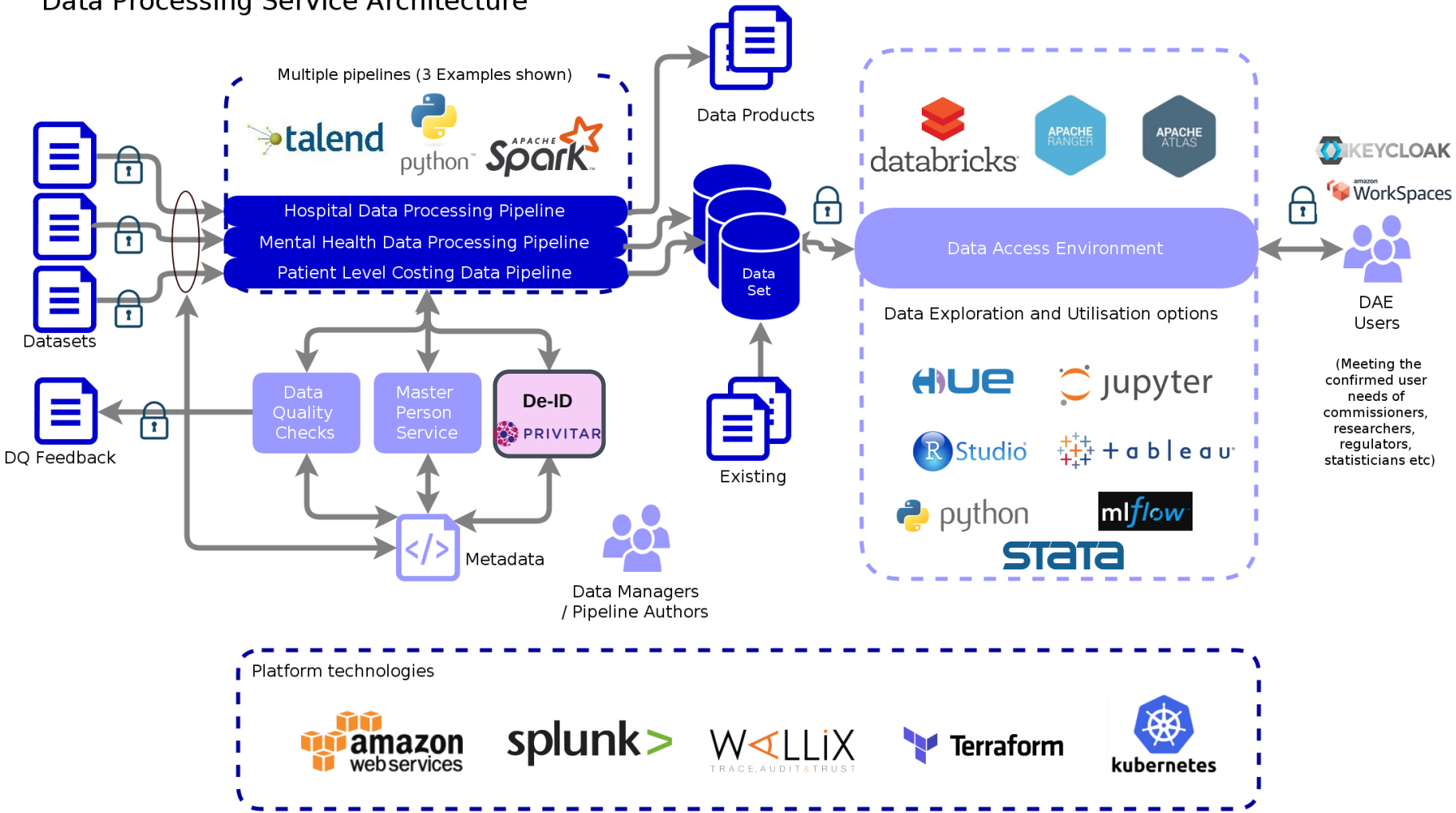
Data Access Environment

- Media
- Researchers
- Charities
- Providers
- Public
- Commissioners
- DH
- Politicians
- Regulators
- ALBs

Our objectives

- **Process NHS data at scale**
  - Support greater collection capacity
  - Faster processing and availability
  - Standardising and simplifying legacy technologies
- **Linking data**
  - Flexible linking of data for richer analytics
- **Enabling flexible, controlled access**
  - Flexible link
  - Minimises the need for data to leave (minimisation)
- **Support data science innovation**
  - Industry recognised analytical tools for deep analysis

Data Processing Service Architecture



DPS Key Deliverables

Name	Metadata Store Phase 1	Virtual Desktop	Talend Integration	Data Managers on platform	Policy Based Access Control
Date The reason for creating the new release	Nov 2020	Dec 2020	Dec 2020	Jan 2020	Feb 2021
Goal	• Consistent data language • Central store for reference and DARS data • Allow collaboration at HDR/NIHR level • Platform automation	• Provide a full and persisting desktop experience for users accessing DAE/TRE	• Enable Data Management to manage data pipelines into DPS to migrate NHSD data and bring in new data	• DMs able to work at pace, independently on the platform	• Move to new Databricks Enterprise 2.0 • Improve speed & specificity of user accesses • Improve access audit capacity
Features The high-level features necessary to meet the goal	• Integrated metadata modelling layer • Single source for ref data • Single source for DARS data for data views and production	• AWS Workspaces • Full multi-screen virtual desktop • Desktop apps (Word/Excel/STATA) • Shared collaboration drives	• Point and Click pipeline generation • Create base data assets for data production • Auto generate Java code	• Independent: ○ Robust Code Promotion ○ Data Curation ○ Data Publications	• Metadata driven data access policies • Policy based Row & Column based minimisations
Metrics The metrics to determine if the goal has been met	• Metastore established Oct 2020 • DARS data incorporated Dec 2020 • Mauro populated and updated Jan 2020	• All existing internal DAE users accessing via virtual desktop by Dec 2020 • DAE/TRE users by Feb 2021	• 1st dataset complete Dec 2020 • 1 further dataset by end Dec 2020 • 7 further datasets by end March 2021	• DSA support reduced 25% by end Nov 2020 • DSA support reduced 50% by end Dec 2020 • DSA support reduced 75% by end Feb 2021	• Policy control to Databricks technically implemented Nov-Dec 2020 • Data access automated by Policy Jan 2021

What these services mean for our customers

- **Channel shift from data extracts to secure access** - reducing the need for data to leave NHS Digital
- **Greater automation** - reducing the need to access identifiable data
- **Moving data onto modern technologies** - reducing risks associated with multiple legacy environments
- **Each data asset is secured independently** - with access restricted according to purpose and governed by DARS
- **Adherence to cloud security good practice** - against the highest data classification with regular cybersecurity testing
- **Data encrypted in transit and at rest**
- **Market leading privacy engineering solution**

Our data supports the national response to COVID-19

Some examples of NHS Digital's work:

- The Recovery Trial
- The Shielded Patient List
- GPES data for pandemic planning and research (COVID-19)
- Test Trace, Contain and Enable
- Dashboards
- Trusted Research Environments

https://digital.nhs.uk/coronavirus

The Recovery Trial

University of Oxford , Recovery Trial (https://www.recoverytrial.net/)

- “Low-cost dexamethasone reduces death by up to one third in hospitalised patients with severe respiratory complications of COVID-19”
- The Recovery Trial used NHS Digital data to support the evaluation of dexamethesone
- Martin Landray, Professor of Medicine and Epidemiology at the Nuffield Department of Population Health, University of Oxford, one of the Chief Investigators from the RECOVERY Trial said: “Since the appearance of COVID-19 six months ago, the search has been on for treatments that can improve survival, particularly in the sickest patients. These preliminary results from the RECOVERY trial are very clear – dexamethasone reduces the risk of death among patients with severe respiratory complications.”

The Shielded Patient List

NHSD SPL Clinical Algorithm

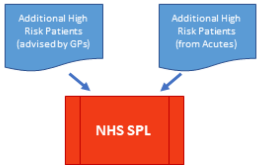
The SPL Clinical Algorithm is developed by NHS Digital at the direction of the CMO for England.

The SPL Clinical Algorithm is published openly at https://digital.nhs.uk/coronavirus/shielded-patient-list/methodology

New versions of the algorithm will be issued if directed by the CMO for England, and published at the above address.

NHS SPL Publication

and additional sources, as directed by NHSE and the CMO for England, currently including:



An updated NHS SPL will be generated by NHS Digital **weekly.** Each list will be based on the latest version of the SPL Clinical Algorithm at that time.

Questions about the NHS SPL Publication process should be directed to SPLQuery@nhs.net

NHS SPL Dissemination

The NHS SPL will be disseminated by NHS Digital to recipients with whom we have a data dissemination agreement in place, which will detail the terms of release and agreements on the intended use of the data, commitments to secure handling and management of the list, and the scope of onwards disseminations.

The NHS SPL contains Patient Identifiable Data and is highly sensitive.

The approved recipients of the NHS SPL are published openly at https://digital.nhs.uk/coronavirus/shielded-patient-list/distribution

The central government SPL service is being run by the Cabinet Office, supported by DWP, Local Government Authorities and others.

*The Senior Manager responsible for the management of the SPL List at NHS Digital is **Mark Reynolds***

Shielded Patient List (SPL)

- With the national availability of CDS and other linkable data we are able to create The Shielded Patient List (SPL) which identifies patients judged to be at the highest risk from COVID-19, based on criteria set by the Chief Medical Officer for England.
- The “at high risk” list size is approximately ~2.2m patients with additions and deductions from GPs and hospitals and minor changes from the national algorithm.
- We are supporting the development of a more complex model for identifying individuals highly vulnerable to COVID-19 by researchers at Oxford.

GPES data for pandemic planning and research (COVID-19)

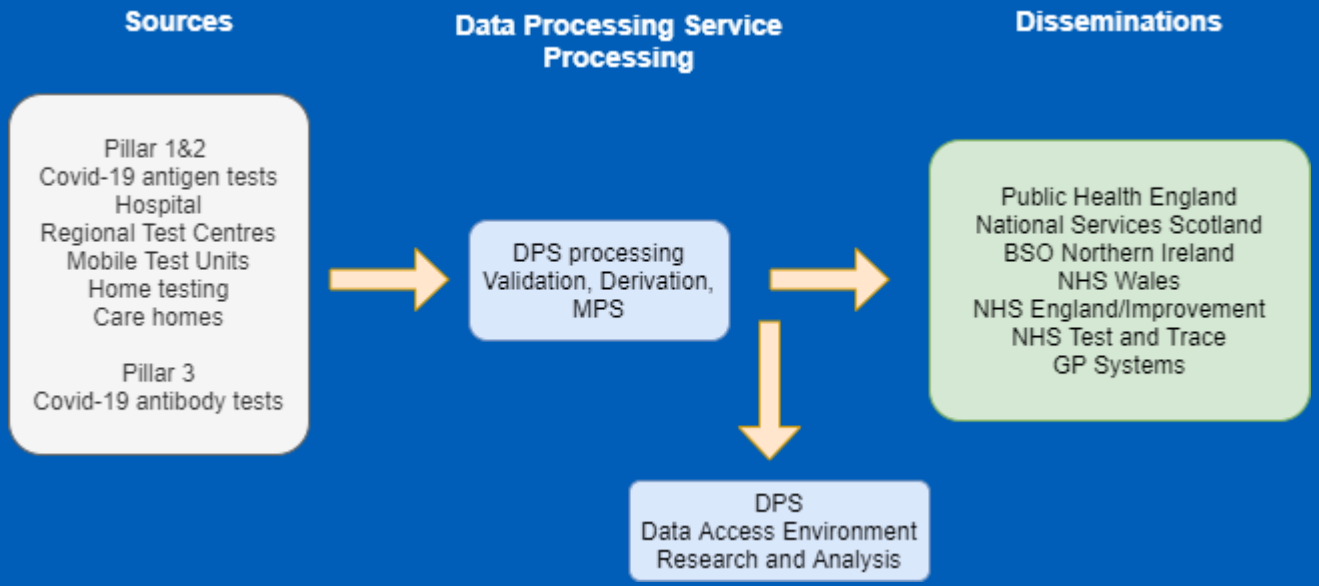
Coronavirus (COVID-19) has led to increased demand on general practices, including an increasing number of requests to provide patient data to inform planning and support vital research on the cause, effects, treatments and outcomes for patients of the virus.

- NHS Digital has been legally directed to collect and analyse healthcare information about patients, including from their GP record, for the duration of the coronavirus emergency period
- This collection was designed to reduce burden on general practices, allowing them to focus on patient care and support the coronavirus response.
- NHS Digital involved the BMA and RCGP in all requests before sharing any data, to consider the views of the profession.
- Data is shared by NHS Digital through established and scrutinised governance through the Data Access Request Service (DARS) and the Independent Group Advising on the Release of Data (IGARD)

Test, Trace, Contain and Enable (TTCE)

This multi-faceted NHSD programme is the digital delivery partner for TTCE, incorporating digital platform for testing, sending test results to GP systems, data dashboards to public and private to LAs, CCGs, PHE and central government.

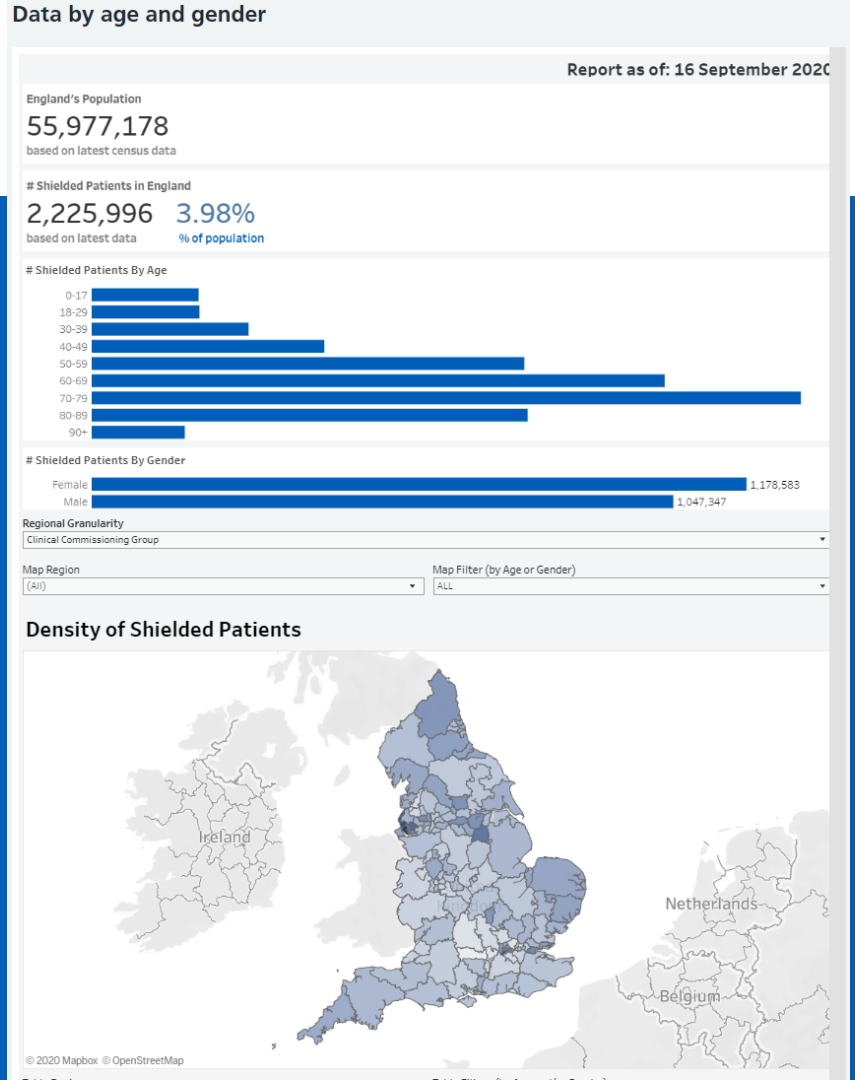
The Data Processing Service receives , processes and disseminates Test data.



Dashboards

As an example here is a part of the Shielded Patient List Dashboard

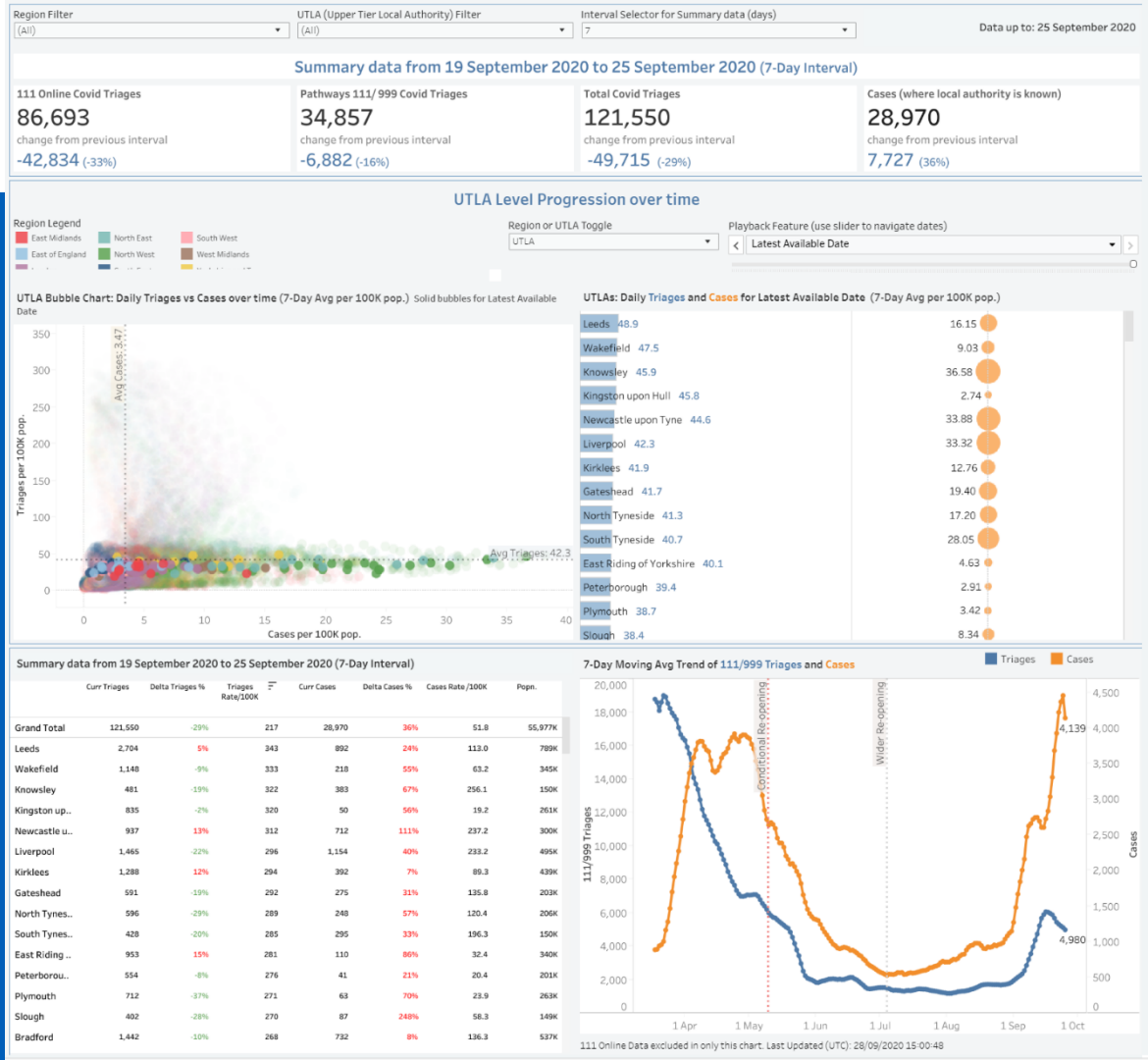
- https://digital.nhs.uk/dashboards/shielded-patient-list-open-data-set
- The dashboards provide a count of living patients that are classified as being on the Shielded Patient List.



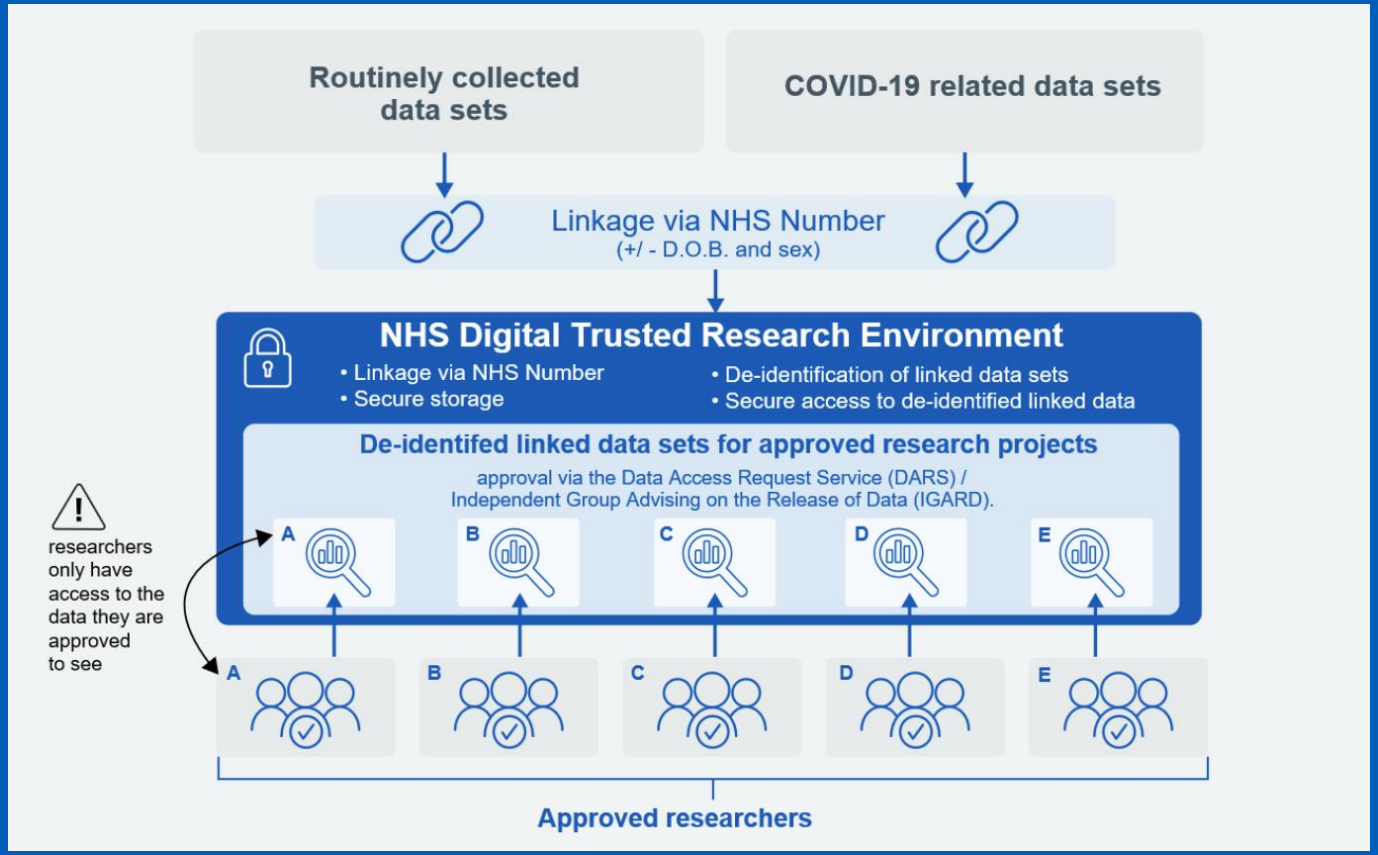
Dashboards

Here is the Shielded Patient List Dashboard

- https://digital.nhs.uk/dashboards/progression
- The dashboard shows both the rate of triage for coronavirus (COVID-19) using NHS Pathways, and the number of people confirmed by a lab to have coronavirus (a positive test).



Trusted Research Environments



NHS Digital

Connect with us

@nhsdigital
company/nhs-digital
www.digital.nhs.uk

Information and technology
for better health and care