DPUK Data Portal & Associated Hubs

Data Provider Handbook

Website : portal.dementiasplatform.uk
Enquires : helpdesk@chi.swan.ac.uk

Dementias Platform UK | Data Portal

Table of Contents

Background	Page 2
Environment	Pages 3-4
Data Deposit Process	Pages 5
Data Access Process	Pages 6
Key contacts	Page 7

Data Portal Overview

Dementias Platform UK Data Portal has been established with MRC funding from the Dementias Platform UK programme of work. Over 10 years the Data Portal has developed robust data management processes with well defined flexible governance approaches and become experts in neurological cohort data. This knowledge and experience has allowed DPUK Data Portal to be able to host Hub spaces for new neurological data collections allowing them to benefit from the mature governance and processes of the DPUK Data Portal whilst retaining their own identity.

This handbook provides information for data providers wishing to deposit data to the data portal for researchers to apply for and access inside a secure Trusted Research Environment (TRE)

Data Portal Overview:

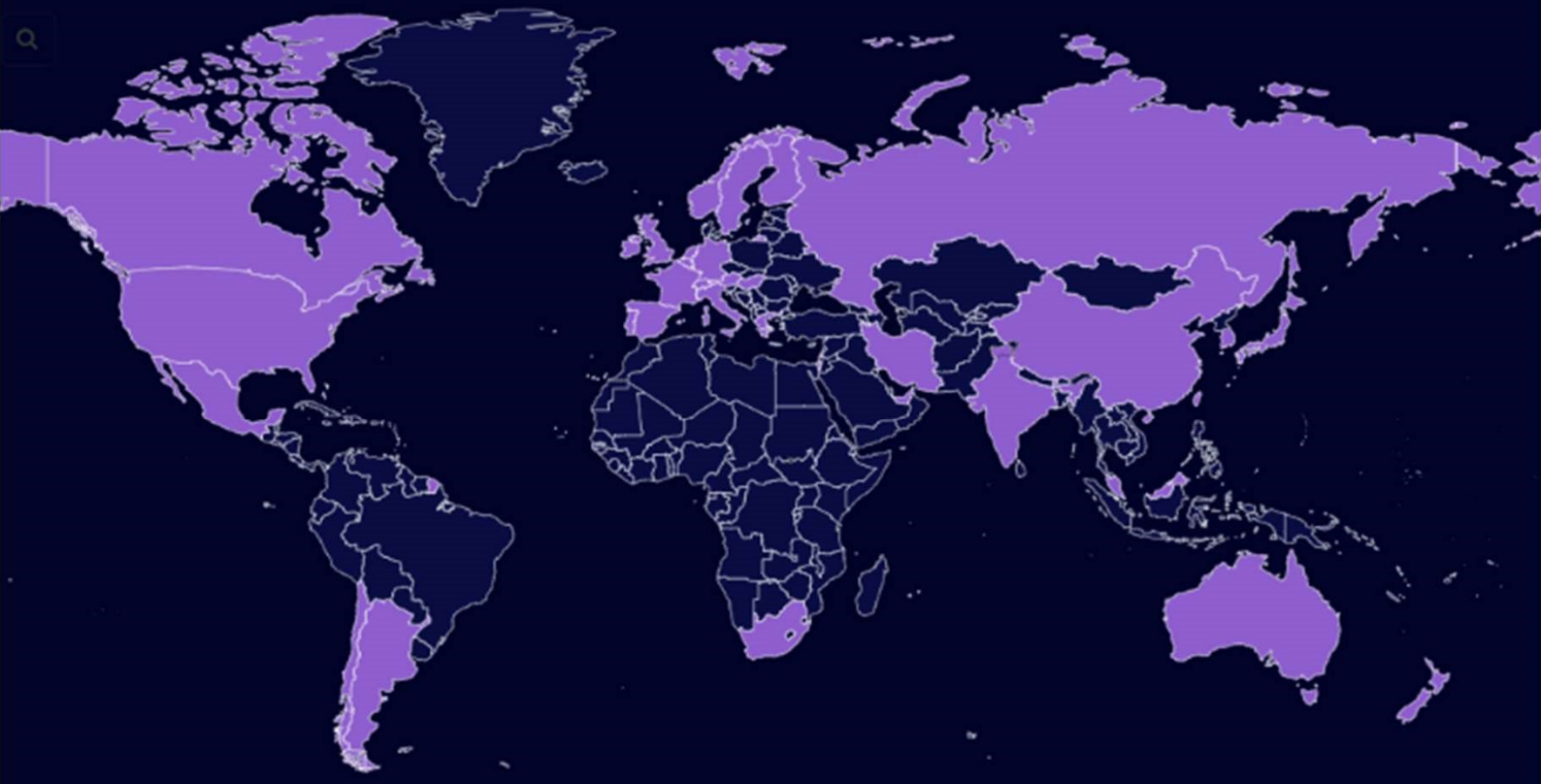
>85 Datasets

>2300 Access Requests

>40 Countries

>560 Project Applications

>1100 Total Researchers



Trusted Research Environments (TRE)

Accessing Data within a Secure Environment

DPUK Data Portal is a secure trusted research environment that incorporates the five safes as defined by Health Data Research UK (HDRUK). These measures and controls along with our ISO27001 and DEA Accredited SeRP infrastructure ensures that the data we hold and provide access to are protected to the highest standards.

Key Features

Data Never Leaves

Approved researchers access the data via a remote virtual desktop - Data never leaves the secure data environment

Analysis Environment

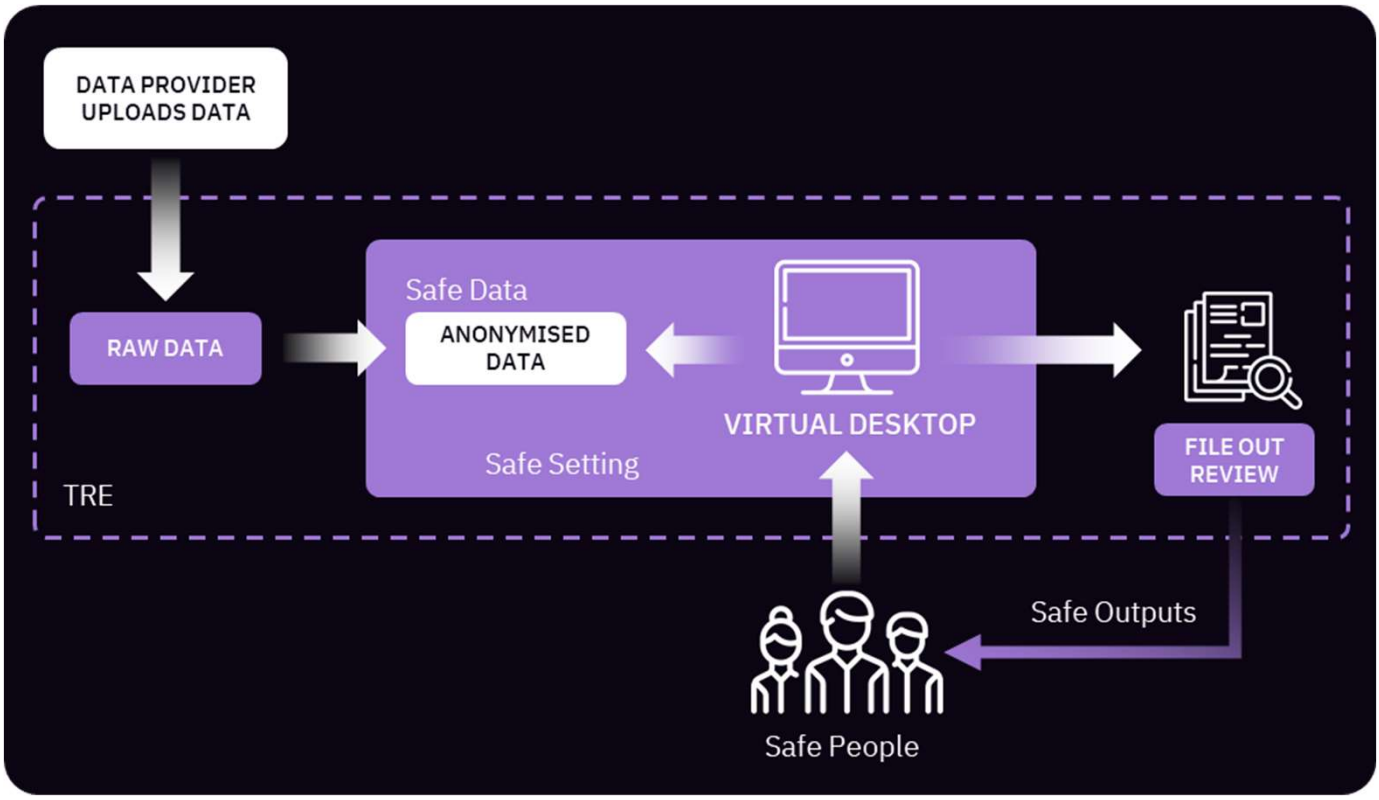
Researchers access specialised virtual desktops preloaded with analysis software

Data Filtering

Automated variable level provisioning ensures only the data approved is accessed

Output Disclosure Checks

All outputs are checked to ensure that they do not contain any row-level data



Trusted Research Environment (TRE)

The Five Safes

Maintaining the security of participants' identities in the data we manage is fundamental to everything we do. DPUK adheres to the Five Safes framework to ensure data is accessible for research while prioritising its safety and security.

1

Safe People

Researchers must be part of a research organisation and each user signs a Data Access Agreement promising to protect the confidentiality of the data at all times which is countersigned by the researchers organisation.

2

Safe Projects

Researchers submit an application for data access to carry out a specific project. Researchers must demonstrate that the use of the data is appropriate and ethical, and that the research benefits the public.

3

Safe Settings

On approval researchers are given access to data requested via a Virtual machine (VM) where they can analyse the data on our secure systems. Data cannot be removed by the researcher from the secure environment.

4

Safe Data

To further protect the identity of participants we de-identify the data by removing names, addresses and any other details that would directly identify data subjects before we make them available for analysis.

5

Safe Outputs

As data cannot be removed from the VM, once work has been completed the results can be requested out of the TRE for publication etc. The results are checked by data analysts to ensure they are safe for release.

Data Deposit Process

The DPUK Data Portal sits within the secure and highly accredited SeRP infrastructure that ensures the data is stored and processed to the highest standards and that data never leaves the platform. The DPUK Data Portal operates an innovative automated application process underpinned by in-built tracking and monitoring. This has allowed DPUK to offer data providers flexible multimodal governance structures that suit individual data providers.

Governance Models

Primary

Data is fully deposited with DPUK with access control passed to the DPUK access committee. This allows DPUK to approve project applications on behalf of the data provider following criteria set by the cohort owner.

Primary - Restricted

Data is deposited with DPUK and data access control remains with the data provider. The data provider reviews data access applications and approval decisions reside with them.

Project Specific

Data is not deposited with DPUK. Each application is reviewed by the data provider and , if approved, a project specific subset of the data is uploaded to the data portal for use by that project - and not for onward sharing.

Internal Only

Data is deposited in a collaborative space within the DPUK Portal. This can only be accessed by the cohort and their collaborators, but includes all of DPUK's software through its virtual environment. It is a good option for cohorts which are not ready to 'onward share' their data, but still want to collaborate with other organisations. It means analysis can be undertaken in a secure environment..

1

Meet the team

Discuss how we can provide a secure environment for onward sharing of your data.

2

Legal agreement

A legal agreement is drawn up and signed by both parties detailing the specifics of the agreement

3

Upload data

An account is set up to enable data to be uploaded to our secure servers

4

Data storage

Data is stored within our SeRP infrastructure which holds full ISO27001 and DEA accreditation

5

Meta data curation

Metadata is categorized, standardized and harmonised to c-surv ontology for data discovery tools

6

Data access

Researchers apply for access and applications are reviewed by triage team and the data providers/SAP

Data Access Process

1

Data Discovery

Researchers use data discovery tools to identify datasets for their research project

2

Application

Researchers complete an online application, detailing the project and submit for review

3

Scientific Triage

The application is checked to ensure that it is complete and the study is scientifically sound

4

Technical Triage

The application is checked to ensure that the project has requested appropriate compute

5

Commercial Triage

For commercial projects, the application is sent to a panel to assess that it is in the public interest

6

AI Triage

AI/ML projects undergo an AI triage to assess the risk of data disclosure on release

7

Approval

The application is sent to the data provider/SAP for review and final decision

8

Data Access

On approval, researchers are provided with a virtual desktop with access to approved datasets

9

Scientific Triage

Datasets are given a DOI and researchers must reference the DOI in any outputs to enable tracking

Key Contacts

Catrin Morris
Operations Co-ordinator
Manages data access applications
Email: catrin.morris@swansea.ac.uk

Elen Golightly
Data Manager
Manages all data management activities
Email: elen.golightly@swansea.ac.uk

Lewis Hotchkiss
Research officer
Manages all neuroimaging activities
Email: l.f.hotchkiss@swansea.ac.uk

Sibil Guntar Vilfan
Data Acquisition Officer
Manages all data acquisition activities
Email: sibil.guntarvilfan@swansea.ac.uk

Professor Sarah Bauermeister
Senior Scientist
Leads data curation activities
Email: sarah.bauermeister@psych.ox.ac.uk

Dementias Platform UK

Powered by SeRP

ISO certified and UK Statistics Authority Accredited

portal.dementiasplatform.uk