Genomics England

Key contact details:

To discuss any aspect of Genomics England dataset you can contact the team via email at info@genomicsengland.co.uk. For specific questions or support with the data access process contact gecip-help@genomicsengland.co.uk.

Genomics England data dictionary:

The Department of Health and Social Care set up Genomics England in 2013 with the aim to transform healthcare, accelerate research and protect citizens. Genomics England data is held within a secure Research Environment and currently contains the following types of de-identified data:

- Genomics data: de-identified genomic sequences, variant and genes from over 100,000 genomes so you can compare mass amounts of data to support your research. You'll also have access to the de-identified genomics data from over 14,700 Covid-19 study participants.
- Clinical data: genomics data on its own can't tell you much without robust patient health data. The Research Environment collates this information so that you can look at the whole picture of what might be causing disease. Adding further data such as proteomic and transcriptomic data and digital histopathology.
- Omics samples.

An overview of the data held within the Research Environment can be found here.

Accessing Genomics England data:

To access the data, researchers must first apply to become a member of the research community, Genomics England Clinical Interpretation Partnership. To join the GECIP your academic institution must be registered and a participation agreement form signed. Further information on the GECIP and the joining process can be found here.

Access and analysis of Genomics England data is managed via a secure, AWS cloud-based Research Environment with patient consent and data protection a priority. To protect patient data, access to the Research Environment is controlled and secure, and only results of analysis can be exported. The secure Research Environment provides approved researchers with a range of open source tools, shared storage, databases and research platforms linking genomic data to a rich set of clinical, phenotypic and longitudinal data. Within the secure Research Environment researchers will be able to bring their own data or containerised applications into private workspaces, and analyse data through Jupyter notebooks.

Further information on the Genomics England secure Research Environment can be found here.

Data Access Pathway:

Researchers will apply for access to Genomics England data via the website. The application pathway looks like this:

Step 1.) Researchers should check whether their institution is registered with GECIP. See below for registering your institution if it is not already registered.

Step 2.) Choose a domain of interest.

Step 3.) Submit an application form. Once researchers have confirmed their institution is registered and have found a GECIP domain of interest they can apply. The application form can be found here.

Step 4.) Log in to the Research Portal. After the application form has been submitted a Research Portal account will be created, where researchers can track the progress of their application. The domain lead will review the submitted application form (this will take approx. 10 working days), the academic institution will validate researcher's affiliation, researcher will be asked to complete online information governance training.

After these steps have been successfully followed and approved, access to the Research Environment will be granted. An independent Access Review Committee examines and approves all requests for access to data before users are granted access to the Research Environment.

User access agreements:

Before an application to Genomics England can be made, the researcher's academic institution must register with GECIP (if not already registered). Further information about registering with GECIP can be found here.

The researcher's institution must also sign a participation agreement form with Genomics England. Only once this has been received will Genomics England allow approved researchers to view or work with the data. Further information and the form to complete can be found here.

Training: Online information governance training is a requirement, only once this has been completed through the Research Portal will access to the data be granted.

Additional information:

To view Genomics England news and explore upcoming events, take a look at this webpage.

End of document.