Edinburgh International Data Facility: data science-enabling infrastructure and services for research and innovation

Bringing together regional, national, and international datasets to create new products, services, and research.

A flexible set of services

The Edinburgh International Data Facility (EIDF) is a set of compute, data and customer services optimised for data analytics and science. EIDF employs a "building block" approach to give it the flexibility required to meet our users' needs.

Businesses, universities, and public sector organisations make use of the EIDF to undertake data-driven research and development.

Customer services

EPCC develops and manages the services provided by EIDF and offers consultancy to help organisations make best use of these services. EPCC provides 9-5 user support during working days through a helpdesk ticketing system. Project owners and users have access to a portal to monitor and manage the services to which they have access.

Data services

EIDF includes a data catalogue where projects can donate research data for long-term storage. Researchers can process these data using EIDF compute services or copy these to another infrastructure. Read access is determined and controlled based on the licence set by the donating party. Data is moved in, moved out, and moved around EIDF through a diverse offering of data transit standards such as S3 and Globus.

The EIDF supports the Data Slipstream project, which is working to bring together the large, diverse and complex datasets vital to Earth Observation (EO) research. This satellite image, which is stored on EIDF as part of the Data Slipstream project, shows a cloud-free mosaic of the Pentland Hills in summer 2018, created on the EIDF from over 100 Sentinel-2 satellite images.

Compute services

EIDF provides three ways for users to interact with its compute and data services: a virtual desktop environment accessible through a web browser, a terminal login, and Jupyter Notebooks. Through these users then interact with a diverse set of compute infrastructure that includes a large and growing set of GPU accelerators.

|epcc|

THE UNIVERSITY of EDINBURGH

EIDF provides customer, data and compute services which are optimised for data science. We offer a range of services from scalable notebooks, through data science desktops to specialist compute platforms. All services are managed and supported by a customer-facing team.

Customer support

- A portal to manage projects and users
- A ticket-based helpdesk during working days
- Consultancy on enabling data science (MLOps)
- Consultancy on performing data science (ML).

Compute elements

- Virtual desktops with up to 64 CPUs and 896 GB memory
- Ultra2 576-core cluster with 18 TB memory
- GPU accelerators with NVIDIA H200, H100 and A100
- Cerebras CS-3 Wafer-Scale Cluster AI accelerator
- Jupyter Notebooks.

The EIDF, operated by EPCC, includes Cerebras CS-3 systems configured as a Wafer-Scale Cluster.

Data transit and storage

- Transit of data via S3, sftp, rsync, Globus, and MFT
- Catalogue of research-ready datasets
- Analytics-ready data on a wide variety of themes
- GitLab to enable shared code and documentation development.

Consultancy

Using a building block approach from the "EIDF toolkit", our engineers can work with you to solve your particular challenges, whether it be data hosting, AI modelling, high-performance analytics – or something entirely new!

Data-Driven Innovation initiative

EIDF is funded through the Data-Driven Innovation (DDI) initiative within the Edinburgh and South East Scotland City Region Deal.

Contact

Please get in touch to discuss how the EIDF can enable your data science projects.
Email us at: eidf@epcc.ed.ac.uk
Apply for a project: https://portal.eidf.ac.uk/

EIDF Edinburgh International Data Facility

THE UNIVERSITY of EDINBURGH | HERIOT WATT

DDI Data-Driven Innovation

Part of the Edinburgh & South East Scotland City Region Deal

https://edinburgh-international-data-facility.ed.ac.uk