



The leading
structural biology
research infrastructure

Integration of Instruct into EOSC Life Sciences Connect Node

Marcus Povey, October 2026



instruct-eric.org



[@instruct-eric.bsky.social](https://instruct-eric.bsky.social)



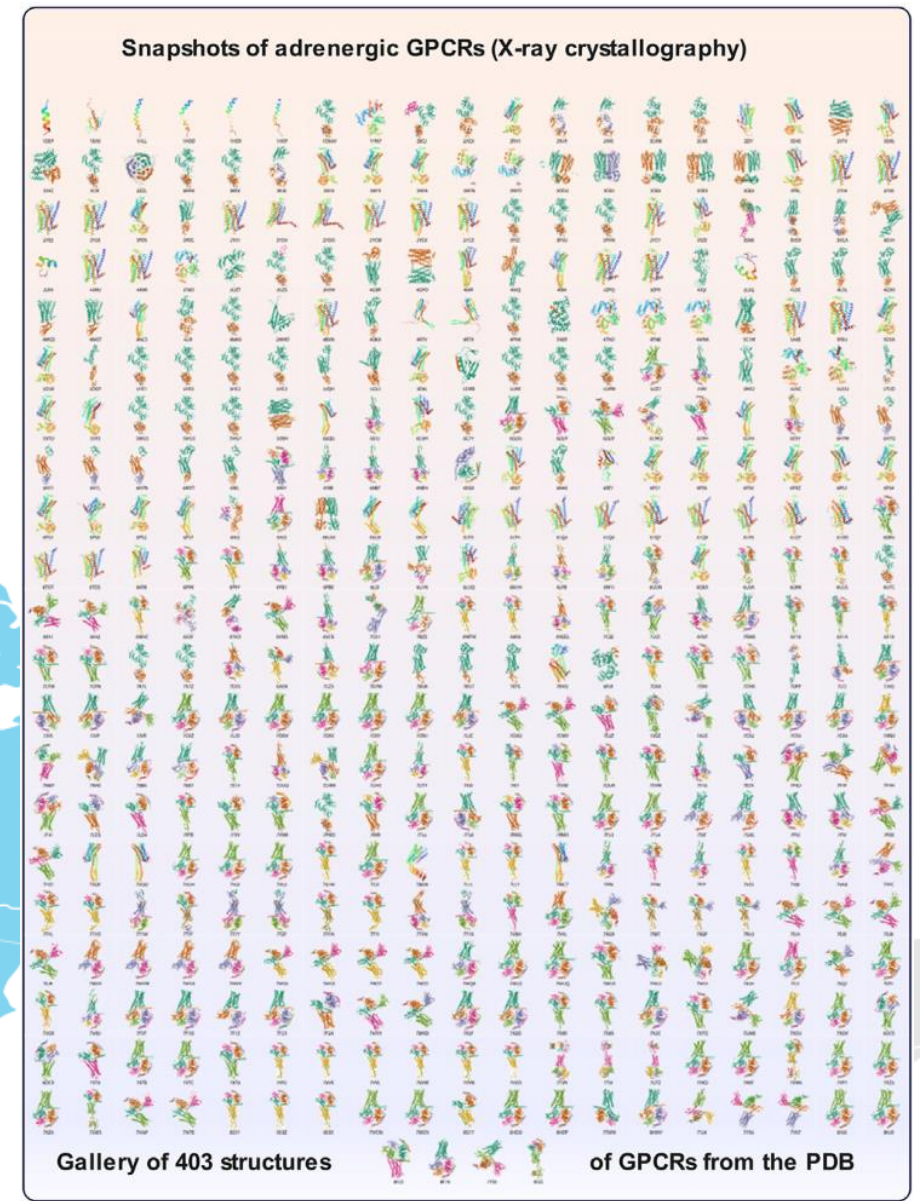
What is Instruct-ERIC

Instruct is a European research infrastructure consortium spread across Europe, with **11 centres, 29 facilities, and over 100 service providers** with expertise in integrated structural biology approaches.

Integrated structural biology combines conventional and novel approaches to revealing the **structure and biophysical properties** of biomolecules.

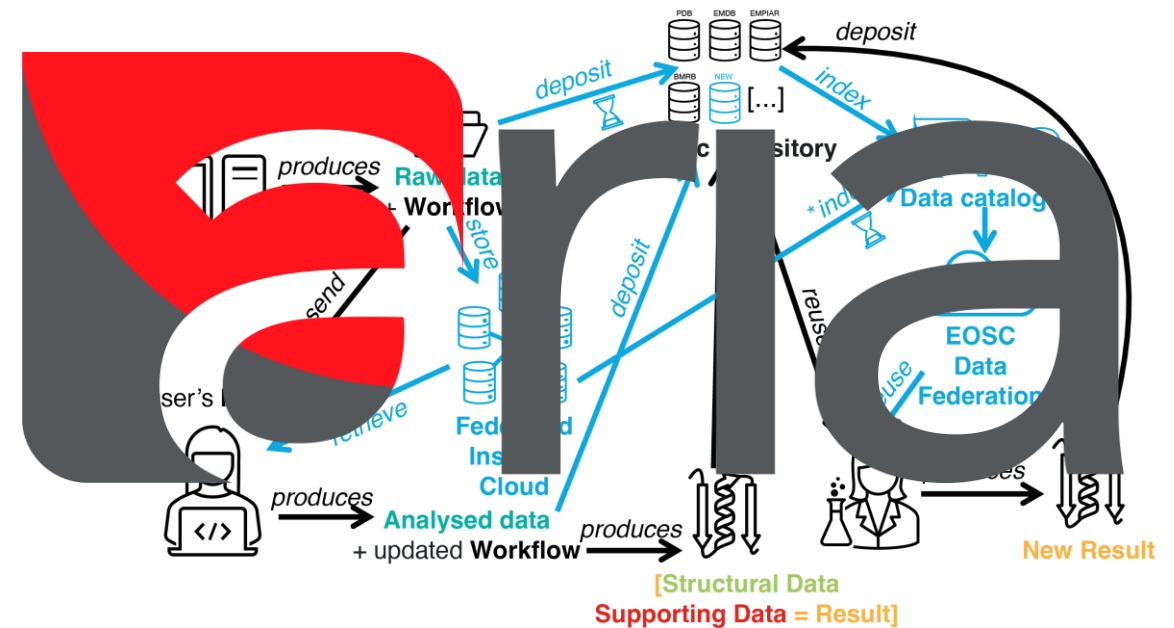
Instruct funds access to facilities for each step of the research lifecycle, from sample prep, preliminary experiments, and high-end data collection.

Centres produce large amounts of raw and processed data and increasingly **raw data, not just processed depositions**, is becoming vital for AI model training and experimental validation



My Work at Instruct: “Cradle to grave” provenance

- Researchers apply through ARIA
- Data, including raw data, is associated with the record in ARIA using (Instruct Fandango lightweight LIMS)
- Result is a “research record”
- All data collected during the steps up to structural characterization are valuable



Sample prep.

Cloning Expression +
Purification
Crystallization, grid
prep. etc.



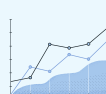
Data Collection

Instrument used
Collection settings



Raw data

Diffraction patterns
Electron micrographs
NMR spectra



Data Processing

Software used
Settings selected
Inputs and outputs



Deposition

3D
structure
deposited





eOSC Node

Life Sciences Connect

Health & Food

Instruct-ERIC

- ARIA and access management
- FandanGO automated raw data/metadata collection/deposition
- Structural data, modelling, analysis, visualisation, lab systems & knowledge tools.



ELIXIR

- Core Data Resources
- 100's of community databases
- National Clouds & Galaxy servers
- LS Login Service

Euro-BioImaging ERIC

- Image analysis as a service
- Image data stewardship service
- BAND virtual desktop
- Community tools for data access and analysis

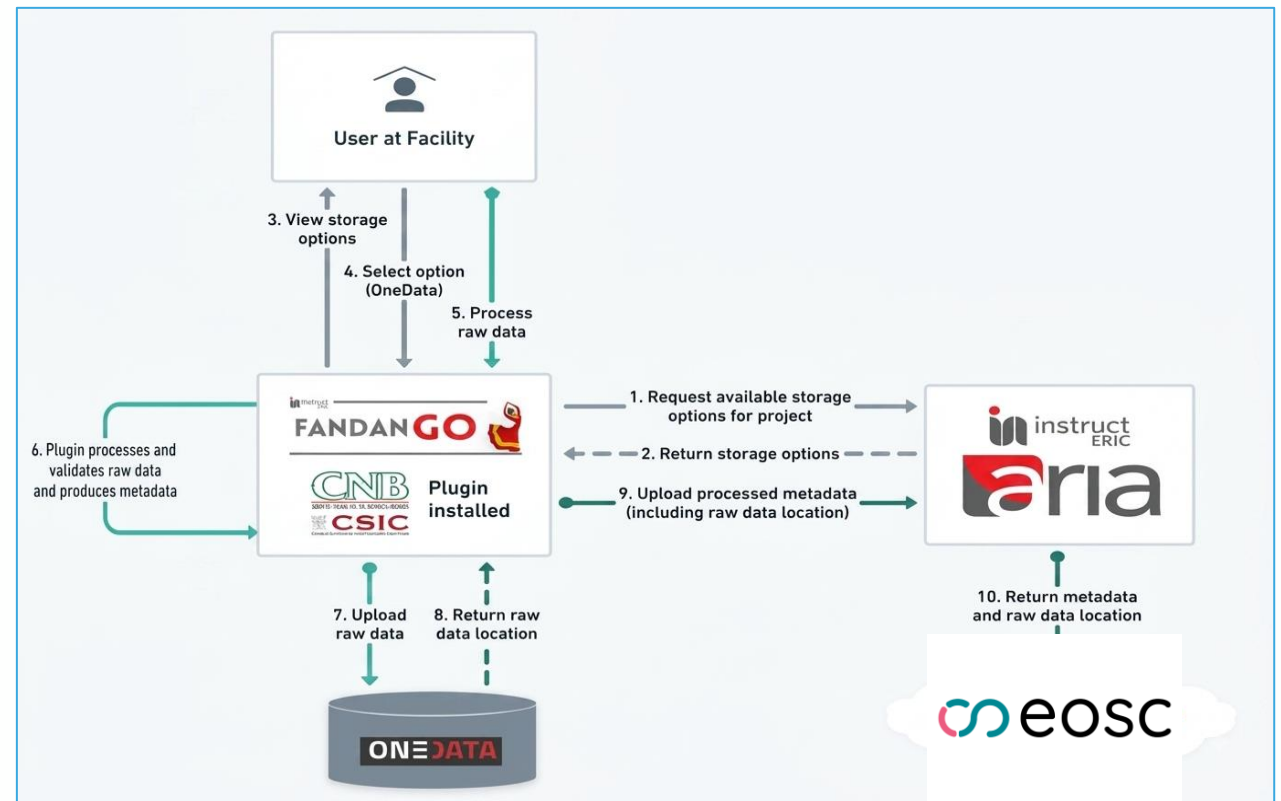


EMBL

- Multi-omics data
- Infrastructure: PDBe, ChEMBL, ENA, MGnify, BIA
- Services: curation, compute and access APIs

Use case: Connecting raw data to EOSC Federation

- Many of our facilities don't have significant long term storage capabilities, but long-term storage of experimental data is necessary to fuel future scientific discovery including by AI.
- We have piloted federated data storage with EOSC-PL & EGI/EUDAT via FandanGO. Connections to national nodes with links to Instruct facilities are under consideration.
- Federated data storage allows facilities to connect via FandanGO, and users to select most appropriate option, whilst also facilitating indexing e.g. via OpenAIRE.





- **Success 1:** Researchers already combine Instruct, ELIXIR, EMBL, and Euro-BioImaging services in everyday work, with a common login (LS Login) now being connected to EOSC
- **Success 2:** A federated approach is viable: storage capacity can be plugged in from across the EOSC community. Our pilot with EOSC-PL, EGI and EUDAT was driven by real centre need
- **Challenge 1:** Four institutions, with their own procedures. Services are often outside our control, so we cannot dictate requirements, and much of the work is in-kind.
- **Challenge 2:** A node needs to offer long-term storage for raw data, not only discovery and access. Volume, funding, and incentives are unresolved

Long-term raw data storage is a gap EOSC could help close.

