## deRSE25 and SE25 Timetables



Contribution ID: 60

Type: Demo

## COPO: A Collaborative Platform for FAIR Metadata in Omics Research

Wednesday 26 February 2025 19:20 (20 minutes)

The Collaborative OPen Omics (COPO) is a data and metadata broker that advances open science by supporting the principles of Findability, Accessibility, Interoperability, and Reuse (FAIR). As reliance on shared data grows, COPO addresses metadata management challenges by using community-sanctioned standards, specifically Darwin Core (DwC) and Minimum Information about any Sequence (MIxS). These standards enable discoverability and reusability of diverse omics data across platforms, which is especially relevant for the life sciences.

COPO streamlines and validates metadata submissions (e.g. samples, reads, assemblies and sequence annotations) through user-friendly interfaces, ensuring consistency and high data quality. Data can be accessed through well-defined application programming interface (API) endpoints, with outputs available in Commaseparated values (CSV), Research Object Crate (RO-Crate) and JavaScript Object Notation (JSON) formats, supporting versatile data use and integration. COPO also safeguards personal data, excluding information such as ORCID identifiers and email addresses from API results, which supports compliance with General Data Protection Regulation (GDPR) and ensures researcher privacy. Leveraging the European Nucleotide Archive (ENA) as a primary data repository, COPO enhances interoperability with databases like BioSample at National Centre for Biotechnology Information (NCBI), fostering scientific collaboration and accessible research findings.

Through the integration of DwC and MIxS standards, COPO enhances metadata structure and context, aiding in data discoverability. The platform incorporates Tree of Life (ToL) projects, enabling users to specify details such as sample locations, collection dates and taxonomy in well-defined spreadsheets or web forms. This information is then mapped to DwC standards for API outputs, ensuring interoperability and consistency. Similarly, MIxS standards can be used ti output the minimum sample information, including environmental context and experimental conditions, ensuring that metadata aligns with community norms.

To improve scalability, reproducibility, and usability, COPO employs modern deployment tools like Docker, whose containerisation enables consistent deployment of its API endpoints and metadata management tools across various environments, reducing complexities in installation and version control.

In summary, COPO represents a substantial advancement in omics data management and dissemination. By adhering to FAIR principles, implementing recognised standards, protecting sensitive information, and utilising advanced technologies, COPO strengthens research interoperability and supports a collaborative open science culture. This platform empowers researchers to document and share findings effectively, advancing biological sciences and facilitating future discoveries.

## I want to participate in the youngRSE prize

no

Primary author: Ms PROVIDENCE, Aaliyah (Earlham Institute)

Co-authors: Ms KU, Debby (Earlham Institute); Dr SHAW, Felix (Earlham Institute)

**Presenters:** Ms PROVIDENCE, Aaliyah (Earlham Institute); Ms KU, Debby (Earlham Institute) **Session Classification:** Poster and Demo Session together with Reception

Track Classification: Policies and Community Building: open source research software