HZB Helmholtz Zentrum Berlin

SEPIA: Integrated Sample Metadata for ICAT Facilities

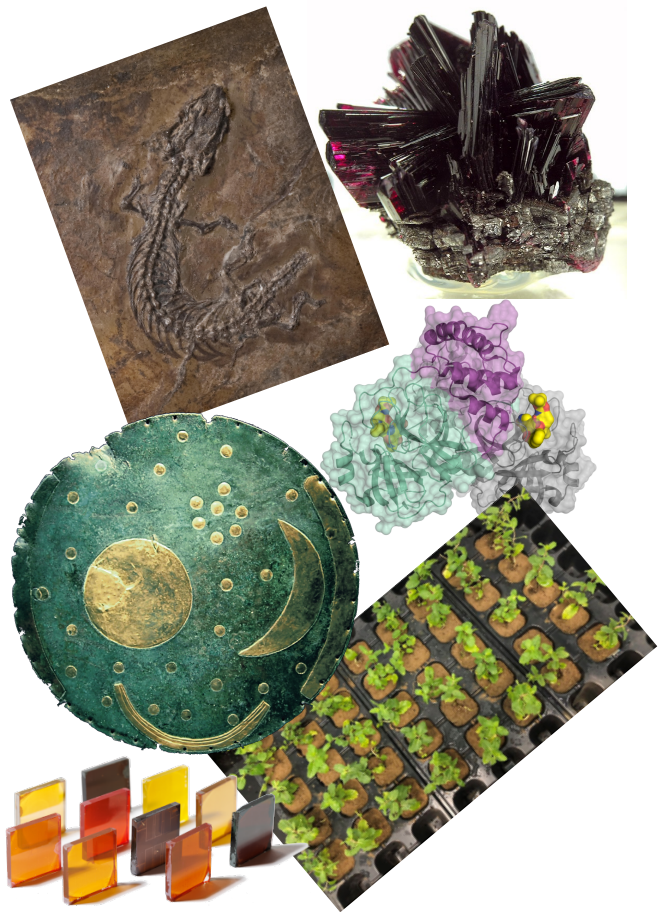
Mojeeb Rahman Sedeqi Katherine Rial Rolf Krahl
Heike Görzig

ICAT Face-to-Face User Meeting — UK, 10 February 2026

CC BY

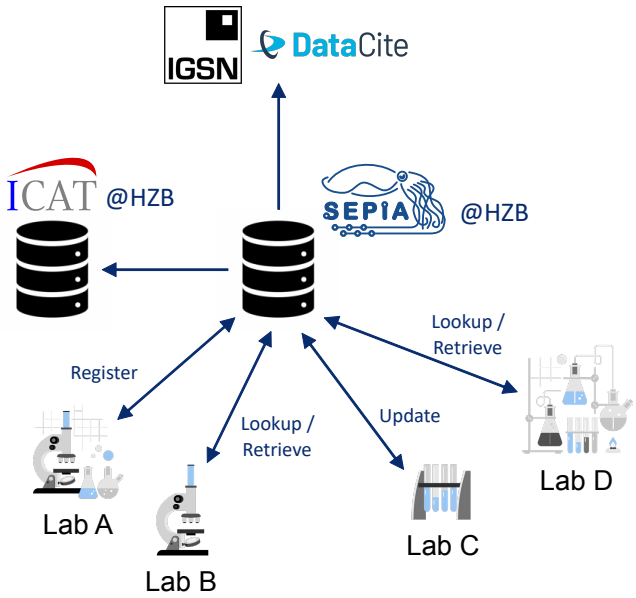
Motivation and Challenges

- **Thousands of samples** per year measured by HZB, with **huge variety**:
  - single crystals, crystal powder,
  - complex structures: solar cells, batteries,
  - macromolecular crystals, e.g. proteins,
  - archaeological artefacts, fossils,
  - artwork, e.g. paintings,
  - living plants,
- Measured at **multiple beamlines**, sometimes **multiple facilities**
- **FAIR** pressure is increasing (from funders & users)



Solution: SEPIA (Sample Essentials, Persistent Identifiers & Attributes)

- **Centralized, persistent, FAIR-compliant sample metadata platform**
- SEPIA tracks **sample history from day zero** using persistent identifiers.
- Provide a **common core metadata schema**, basically citation metadata.
- Allow **sample type specific schema** extensions.
- Provide an **API for the integration** into lab workflows.
- Allow linking all **documents, lab notebooks, datasets** related to the sample with its id.

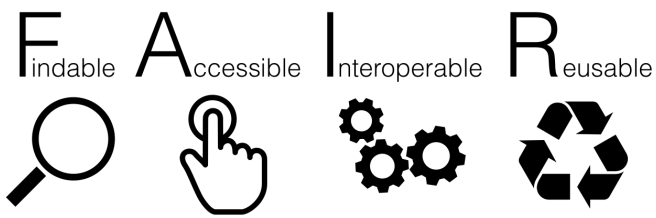


Solution: SEPIA ...

- **Global PIDs** via IGSN + DataCite
- Provides the **IGSN landing page**.
- **Tracks**:
  - provenance
  - modifications
  - links to datasets & experiments
- Designed with **close integration** concern with **HZB metadata catalogue ICAT**.
  - **SEPIA = ICAT Module for sample** (reference via PID)
- Status: **SEPIA is in development**.

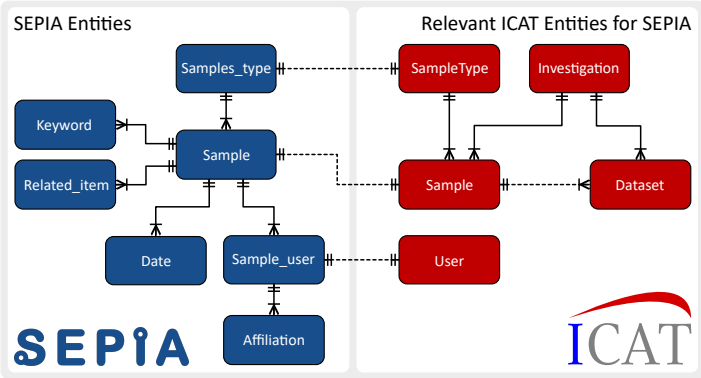
SEPIA and FAIR Sample

- **Findable:** persistent identifiers for stable referencing.
- **Accessible:** landing pages + REST API.
- **Interoperable:** standard schemas - DataCite, & OpenAPI specification.
- **Reusable:** full provenance, sample-type specific metadata.



Architecture & ICAT integration

- **SEPIA runs as a separate service**
  - PostgreSQL backend
  - Flask REST API (OpenAPI)
  - Next.js web frontend
- **Tight ICAT integration**
  - SEPIA backend connects to ICAT via `python-icat`
  - Authentication and authorization are fully handled by ICAT
  - ICAT `sessionId` is used consistently across SEPIA



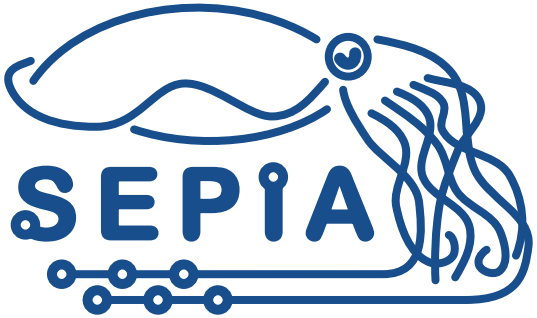
Architecture & ICAT integration ...

- **ICAT remains responsible for**
  - proposals
  - investigations
  - datasets
- **SEPIA provides**
  - sample registration
  - PID minting (IGSN / DataCite)
  - sample landing pages
- **Clean separation of concerns**
  - ICAT stores only the **sample PID reference**
  - SEPIA manages the full, FAIR sample metadata

Live demo — what I will show (5 min)

- **Register a new sample** (SEPIA UI), and Already registered sample **via IGSN**
- **Enrichment** of sample metadata
- **List and search** samples
- **List linked samples** with an investigation
- Show **sample landing page**
- Quick **OpenAPI** / curl example for automation

Source Code



Acknowledgment

- SEPIA is implemented with contribution from the Helmholtz Metadata Collaboration (HMC).
- It is developed in close cooperation with
  - DAPHNE4NFDI and
  - the ROCK-IT project.