KU LEUVEN

Data Management at the FIBEr Lab

ManGO User Day 2024

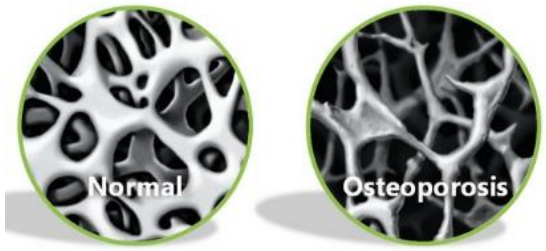
Heleen Fehervary & Ronny Moreas

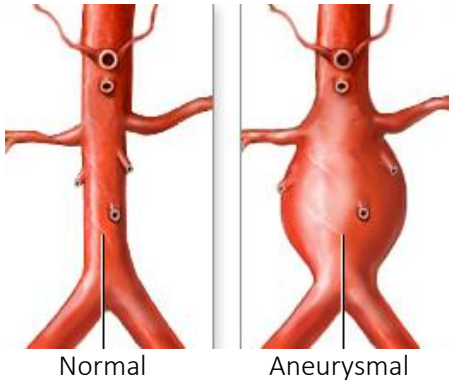
FIBER

We test what's vital for you

WHY do we need biomechanical experiments?

Understanding mechanics-related pathologies





Development of treatments, medical products, diagnostic tools

WHAT are biomechanical experiments?

Material characterization

Aorta

Medical device testing

Fracture plate

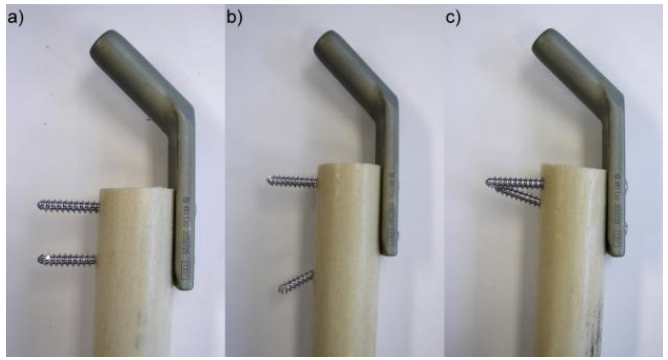
Non-exhaustive list of the **mechanical properties** we can measure:

- compliance or **stiffness** (e.g. Young's modulus),
- **strength** (tensile, compressive, flexural, peel, tear) and elongation,
- energy to **failure** and fracture toughness,
- Poisson's ratio,
- crack growth velocity,
- **fatigue** & fatigue life,
- visco-elastic properties (creep, stress relaxation)
- …

WHEN are biomechanical experiments useful?

at multiple stages of product development

R&D stage



Regulatory

INTERNATIONAL STANDARD **ISO 7198**

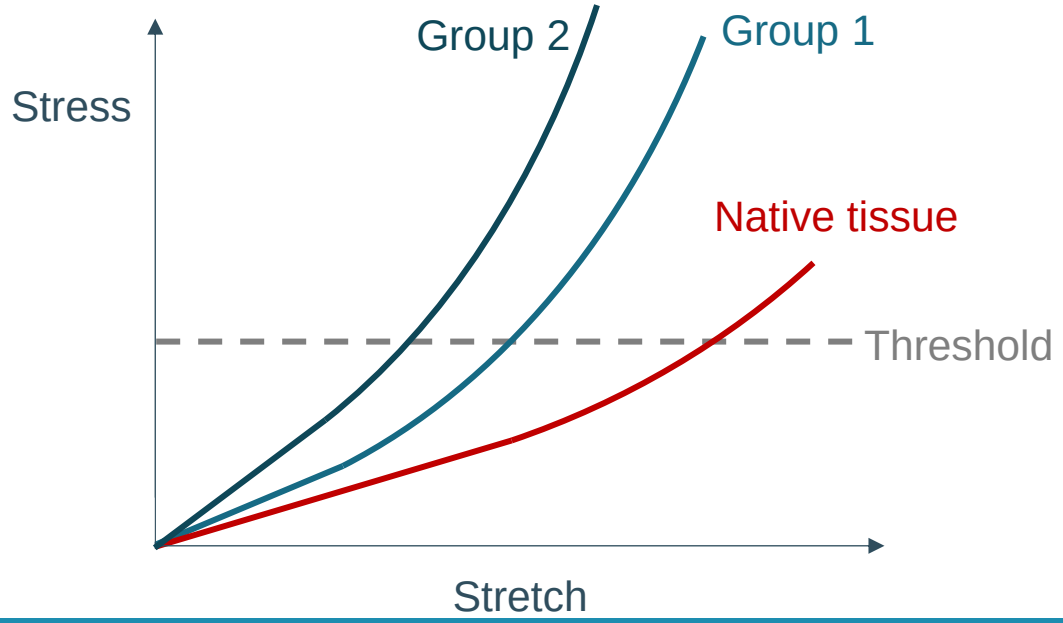
Cardiovascular implants and extracorporeal systems — Vascular prostheses — Tubular vascular grafts and vascular patches

Quality control

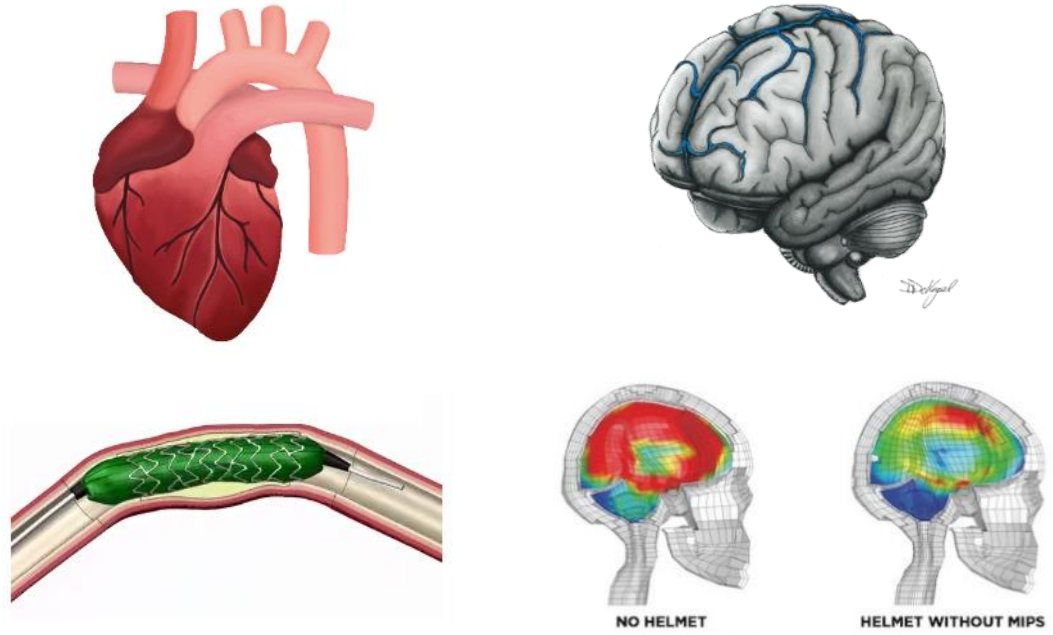
HOW can data from biomechanical experiments be used?

Direct comparison

- Strength
- Stiffness/Elasticity
- Visco-elasticity
- Permeability
- Fracture behavior
- …



Input for in silico models



Grogan et al. 2015, Kleiven et al. 2006

What is important for good data management at FIBEr?

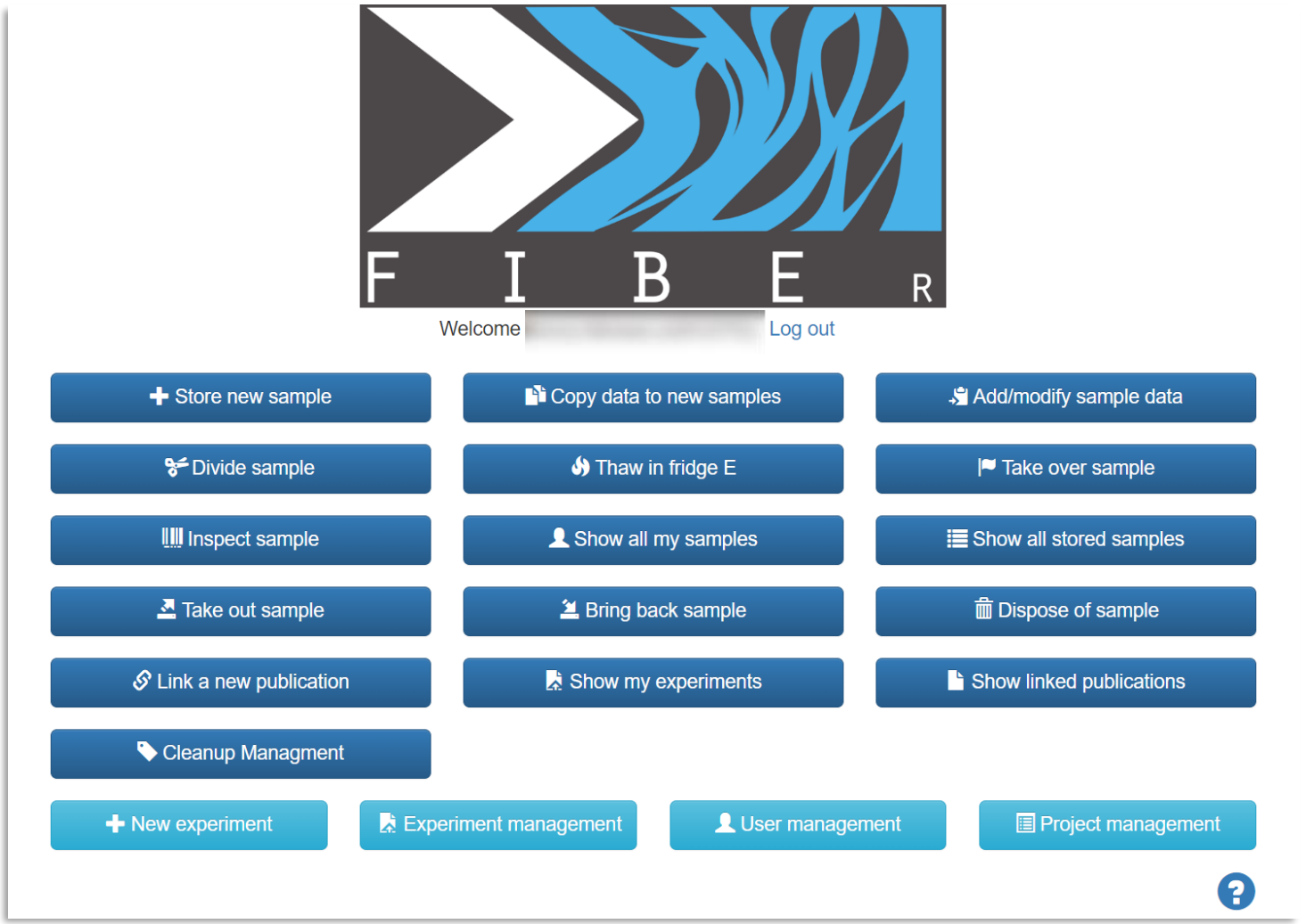
- Different types of clients
  - Lab users (internal, external)
  - FIBEr team running service projects
- Samples (storage, traceability, linked data, …)
- Large portfolio of testing equipment
  - Almost all measurements are digitalized
  - Some data is combined in post processing
  - Different data sizes (few KB to a few hundreds of GB)
- Sharing raw/processed data within/outside KU Leuven
- QMS: back-up, confidentiality, data integrity

Overview Data Management at FIBEr

1. Sample, project and experiment (meta)data
2. Test data acquisition
3. Data ingestion into ManGO
4. Move completed data ingests to project collection
5. Data processing
6. Archival of completed projects

Sample, project and experiment (meta)data

- **FIBEr Sample Manager**: Webapp for management of samples, projects and experiments

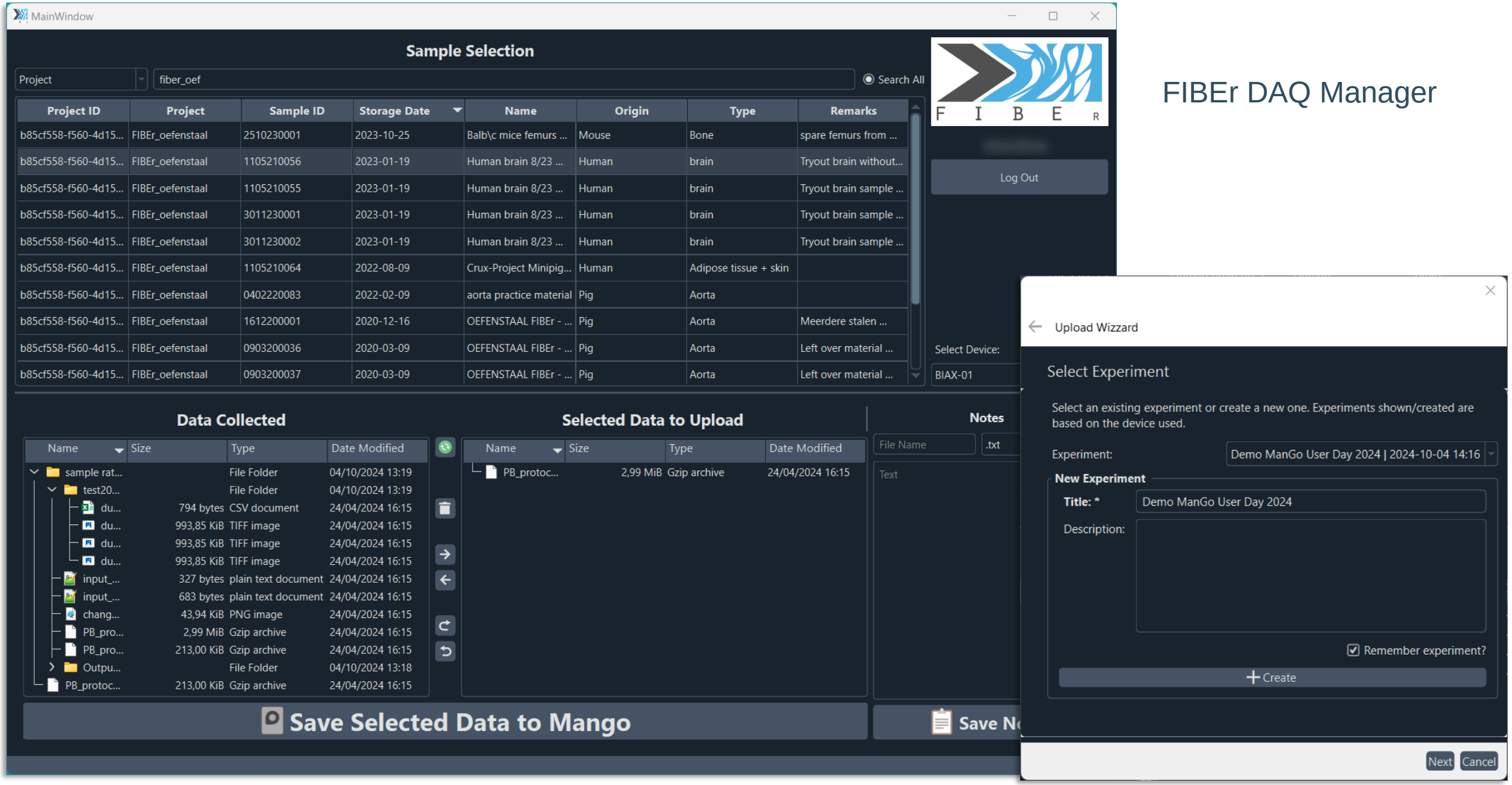


Test Data Acquisition

- **FIBEr DAQ Manager App**
  - Installed on instrument computers
  - Collect data and metadata to be added to dataset for upload
- After a test run the operator:
  - Selects files to be added to dataset
  - Adds metadata (from FIBEr Dashboard)
    - Instrument
    - Sample
    - Project
    - Experiment
    - Test Operator
- Dataset moved to drop location for upload to ManGO

1105210056-2410041417

FIBEr DAQ Manager



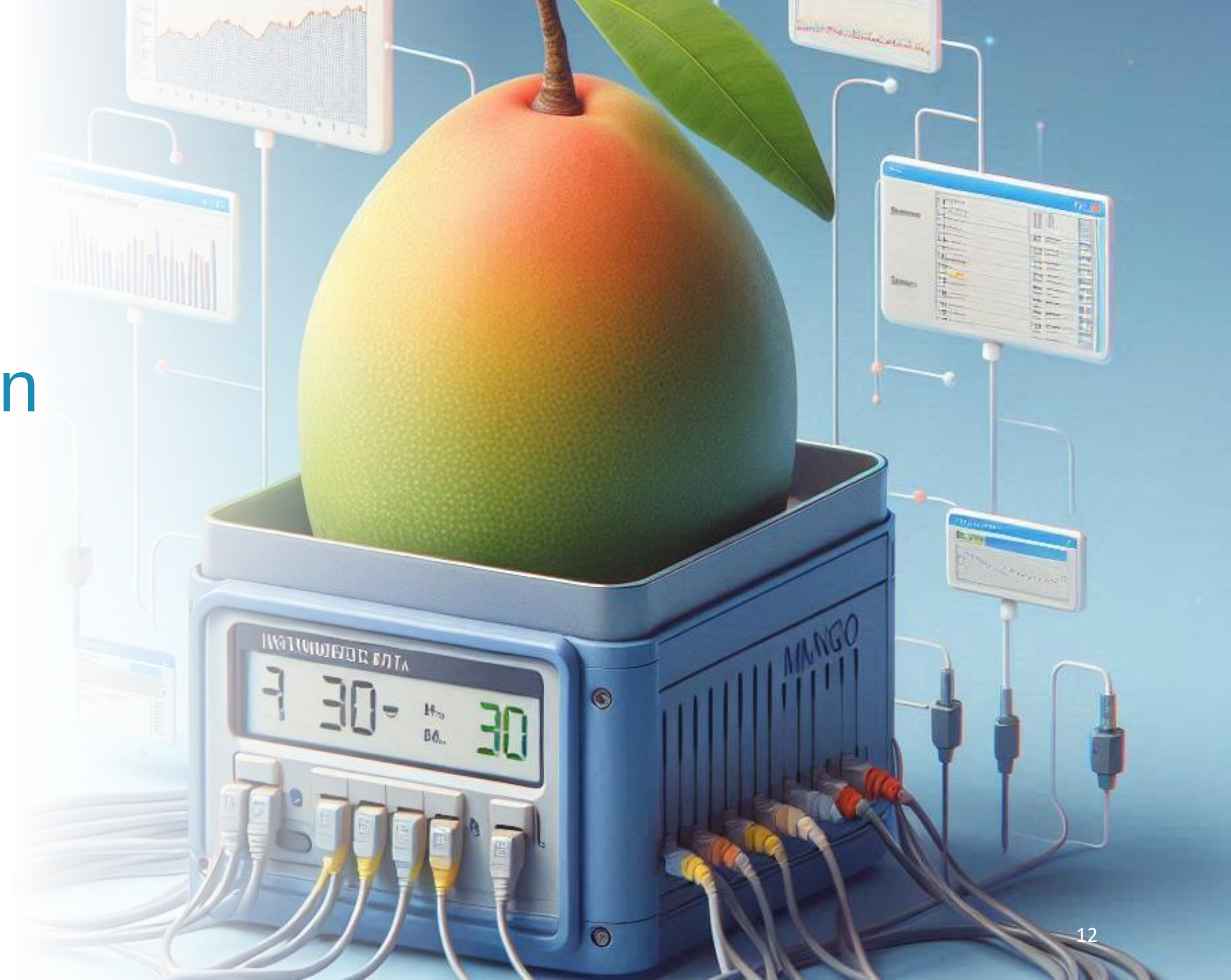
Test Data Acquisition

- Metadata file (json) added to dataset for upload
- Dataset with metadata moved to drop location for upload to ManGO

1105210056-2410041417


```
{
    "acquisition_datetime": "2024-10-04T14:17:47",
    "device": {
        "model": "Biaxial Testing Machine, Messphysik - Zwick/Roell",
        "name": "BIAX-01"
    },
    "experiment": {
        "id": "5f7c7fdda75a8711261e4ea7",
        "name": "Demo ManGO User Day 2024"
    },
    "operator": {
        "name": [redacted],
        "surname": [redacted],
        "uid": [redacted]
    },
    "project": {
        "id": "b85cf558-f560-4d15-a72d-36e1fbce1ed5",
        "name": "FIBEr_oefenstaal"
    },
    "sample": {
        "id": "1105210056",
        "name": "Human brain 8/23 Exp3.1"
    }
}
```

Data Ingestion into ManGO



ManGO Ingestor

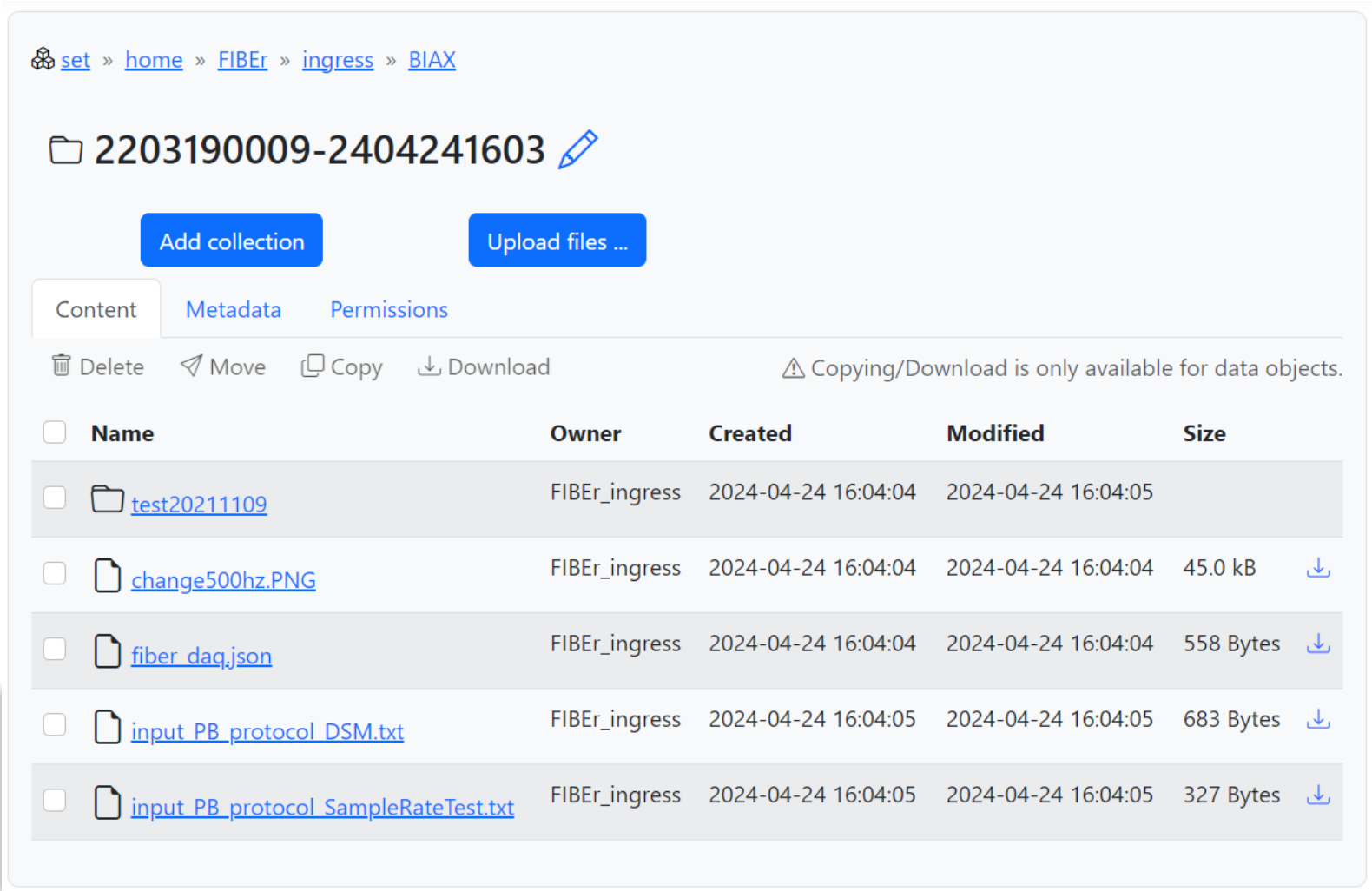
- Linux-based device designed to facilitate the **ingestion of instrument data and metadata from multiple devices** into the ManGO active data platform.
- Purpose
  - Streamline the process of ingesting and managing scientific data **across different platforms and environments with central control**.
  - Avoid installation of software on instrument computers by **pulling data directly from SMB/CIFS (or NFS) share(s) defined on the instrument computers**.
  - **Extendable** with custom handlers for metadata extraction

ManGO Ingestor

- Foundation
  - **Built upon the iRODS Capability Automated Ingest (iCAI)** to provide a library of common ingestion handlers with enhanced configuration features.
  - Includes base **handler classes** for different types of data ingestion, such as RunFolderIngestHandler, BCLIngestHandler, and FASTQIngestHandler.
    - Inspired by bihealth/rodeos-ingest: Code for ingesting omics data into omics storage based on iRODS capabilities (github.com)
  - Cloud-based deployment of software components
    - K3s lightweight Kubernetes distro great for edge deployments in labs
    - ManGO Ingest Docker image
    - ManGO Ingest Helm chart to facilitate configuration, deployment and updating of all software components

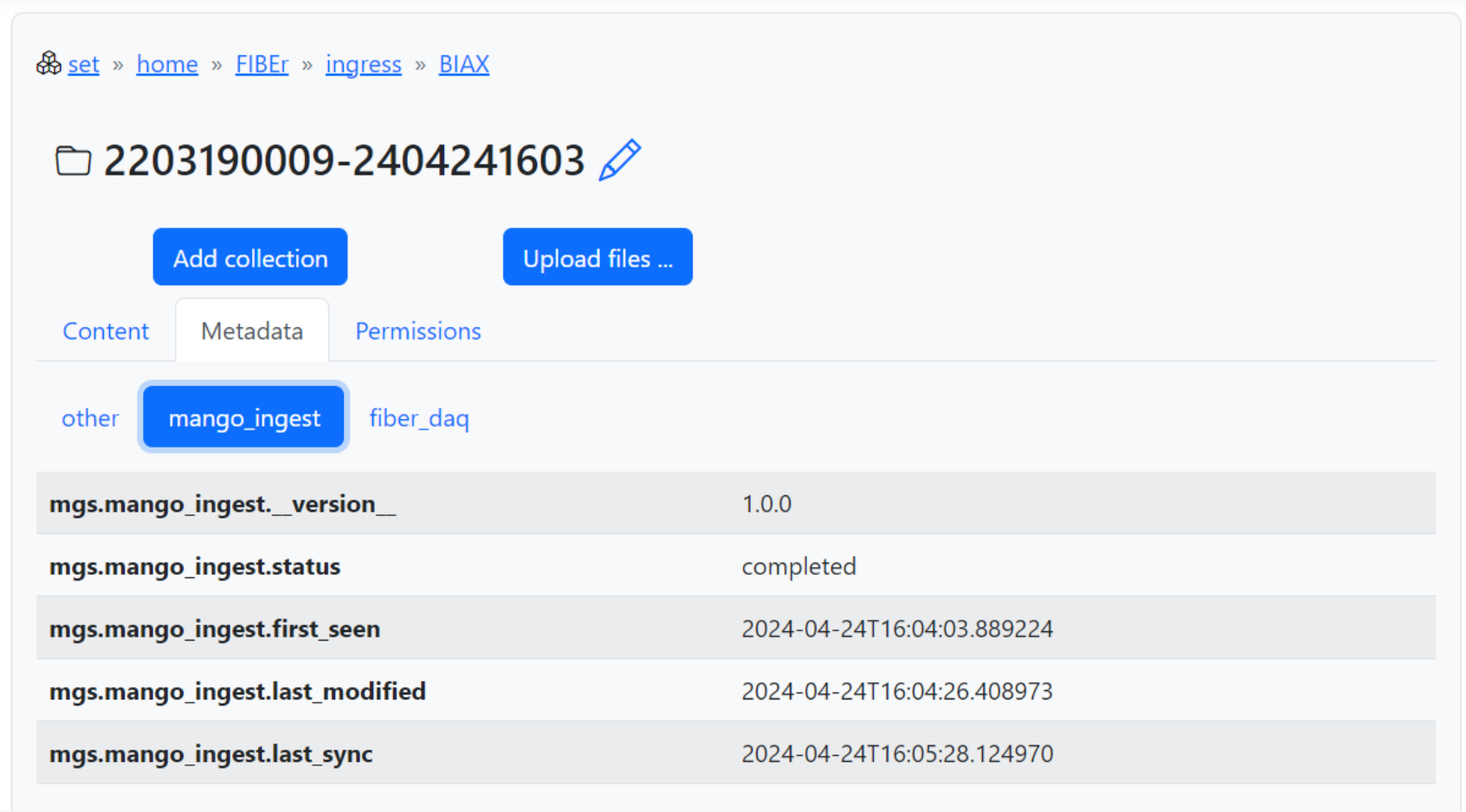
Data Ingestion into ManGO (ingress collections)

- The ManGO Ingestor syncs datasets from the DAQ device drop locations to an **ingress collection per source device**



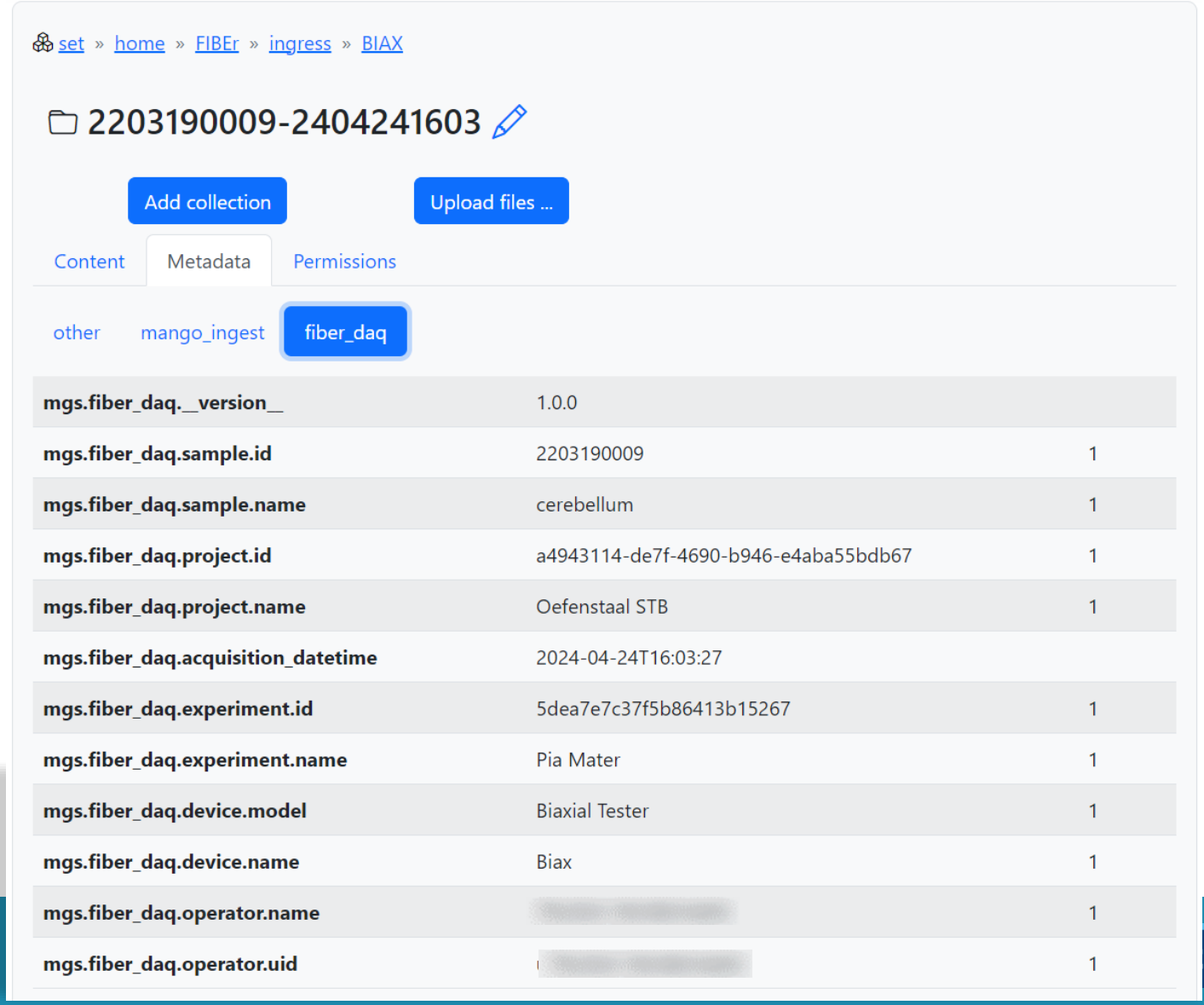
Data Ingestion into ManGO (ingest metadata)

- ManGO Ingestor adds metadata to datasets about the sync status
- Ingestor uses customizable hook method to detect completion of a test/measurement sequence



Data Ingestion into ManGO (other metadata)

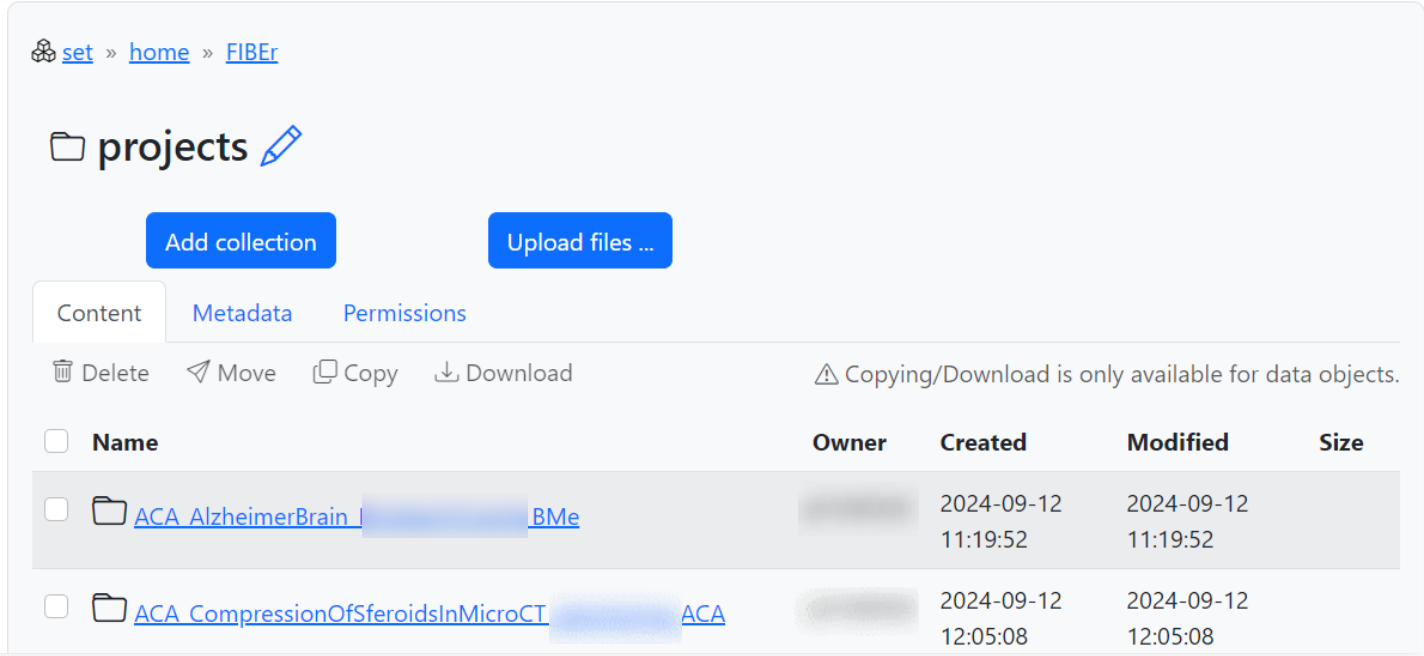
- Acquisition metadata is added to the dataset collection
- ManGO Ingestor can be extended with custom handlers to extract and add metadata for specific instruments

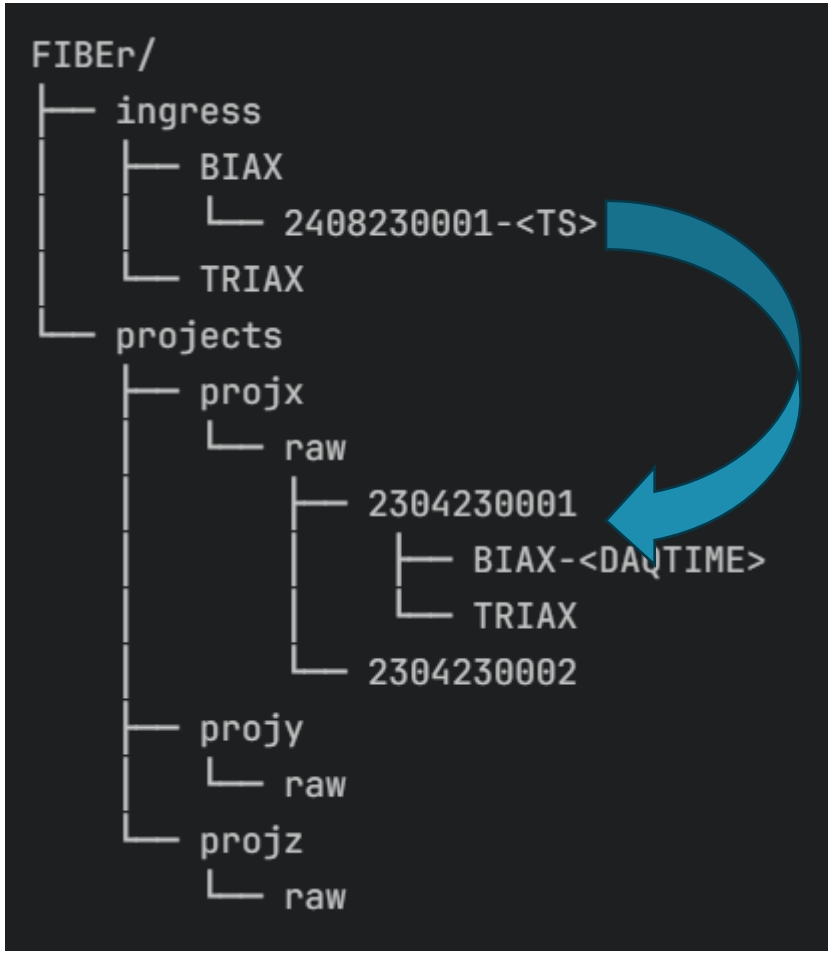


mgs.fiber_daq.__version__	1.0.0	
mgs.fiber_daq.sample.id	2203190009	1
mgs.fiber_daq.sample.name	cerebellum	1
mgs.fiber_daq.project.id	a4943114-de7f-4690-b946-e4aba55bdb67	1
mgs.fiber_daq.project.name	Oefenstaal STB	1
mgs.fiber_daq.acquisition_datetime	2024-04-24T16:03:27	
mgs.fiber_daq.experiment.id	5dea7e7c37f5b86413b15267	1
mgs.fiber_daq.experiment.name	Pia Mater	1
mgs.fiber_daq.device.model	Biaxial Tester	1
mgs.fiber_daq.device.name	Biax	1
mgs.fiber_daq.operator.name	[redacted]	1
mgs.fiber_daq.operator.uid	[redacted]	1

Move Datasets to Project Collections

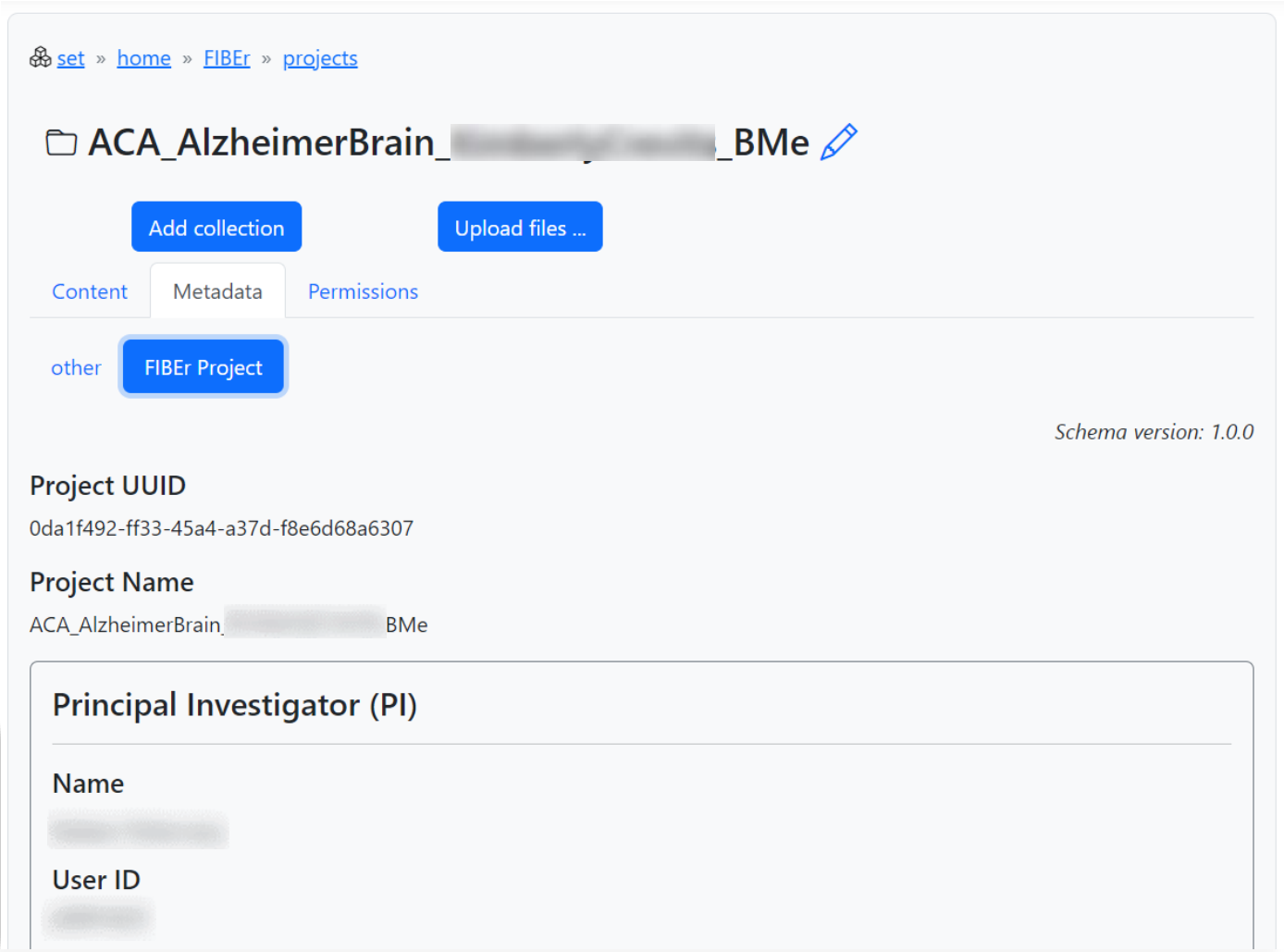
- **ManGO Flow** moves completed data ingests from ingress collection to raw collection of the associated project





Move Datasets to Project Collections (project metadata)

- Target project collection is determined by project metadata set on project collections and the acquisition metadata set on the ingested dataset



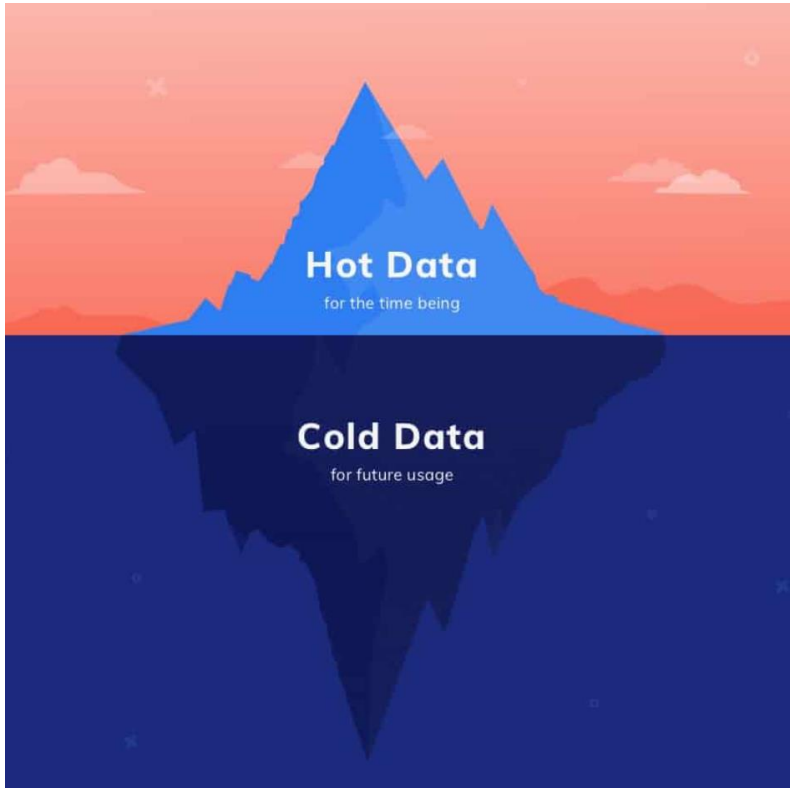
Data Processing

- Workflow depends on project and data types
- Generic workflow using Globus for data transfers
  - Transfer raw data to workstation or HPC for processing
  - Upload of results
  - Still early-stage ideas to be worked out

globus

Archival of completed projects (in the future)

- Move data of completed projects to **cold storage**
  - Long-term storage
  - *Qualified* and well described *immutable datasets*
  - May be re-used or re-examined in the future
    - Enough *metadata* provided, to allow the datasets to be re-used
    - *Searchable* in the long-term



Any Questions?

References

- https://fiber.biomech.be
- https://gitlab.kuleuven.be/setit/rdm/mango-ingest
- https://gitlab.kuleuven.be/setit/rdm/mango-ingest-helm-chart
- https://github.com/irods/irods_capability_automated_ingest
- https://github.com/kuleuven/mango-mdschema
- https://rdm-docs.icts.kuleuven.be/mango/
- https://rdm-docs.icts.kuleuven.be/globus/