

Handling Data for the Ernst Ruska-Centre

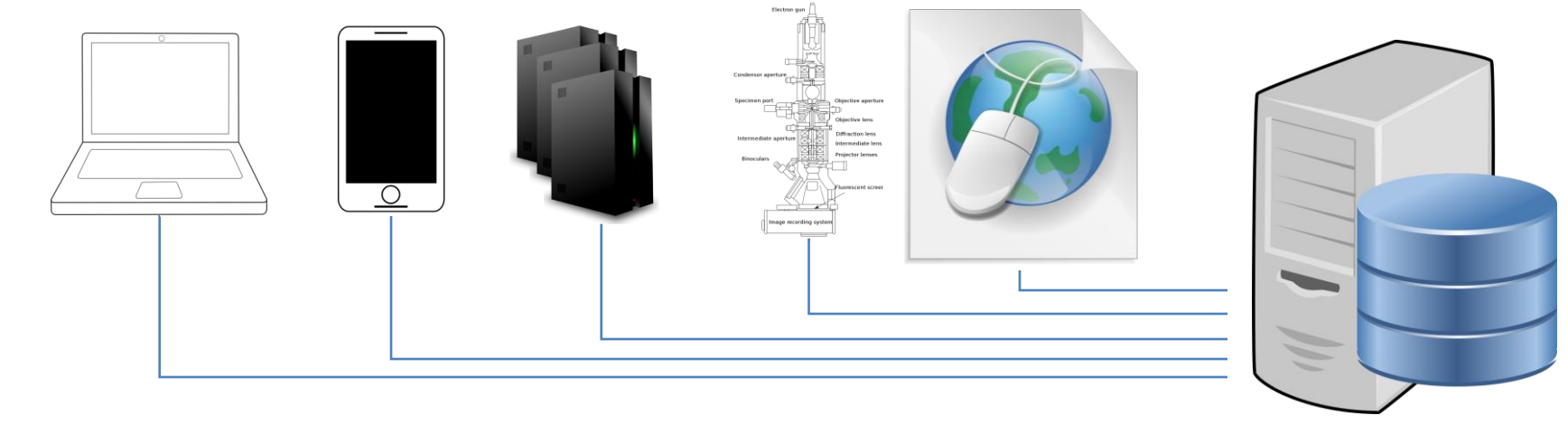
Dieter Weber^{1*}, Alexander Clausen¹, Rafal E. Dunin-Borkowski¹

¹ Ernst Ruska-Centre, Forschungszentrum Jülich, 52425 Jülich, Germany *d.weber@fz-juelich.de



The Ernst Ruska-Centre (ER-C) for Microscopy and Spectroscopy with Electrons is a national user facility open to universities, research institutions and research laboratories in industry.

The ER-C is run conjointly by Forschungszentrum Jülich GmbH and RWTH Aachen University. The facility is located on the campus of Forschungszentrum Jülich GmbH, a major interdisciplinary research centre and member of the Helmholtz Association.



Landing page er-c-data.fz-juelich.de

ER-C Data Management

About
This system stores data for the Ernst Ruska-Centre at Forschungszentrum Jülich. It can handle large scale data with sufficient performance and allows access from microscopes, PCs, processing workstations and mobile devices for institute members and collaborators.

Management
Manage projects, sessions, ingest and access permissions (OpenIDConnect)

Nextcloud
Synchronize, collaborate, ingest data via Nextcloud

Network Filesystem
Directly access files from your favorite programs via network filesystems (OpenIDConnect)

Jupyter Notebooks
Analyze and process data with Jupyter notebooks (OpenIDConnect)

File Transfer
Use standard tools like scp or rsync to transfer data in or out

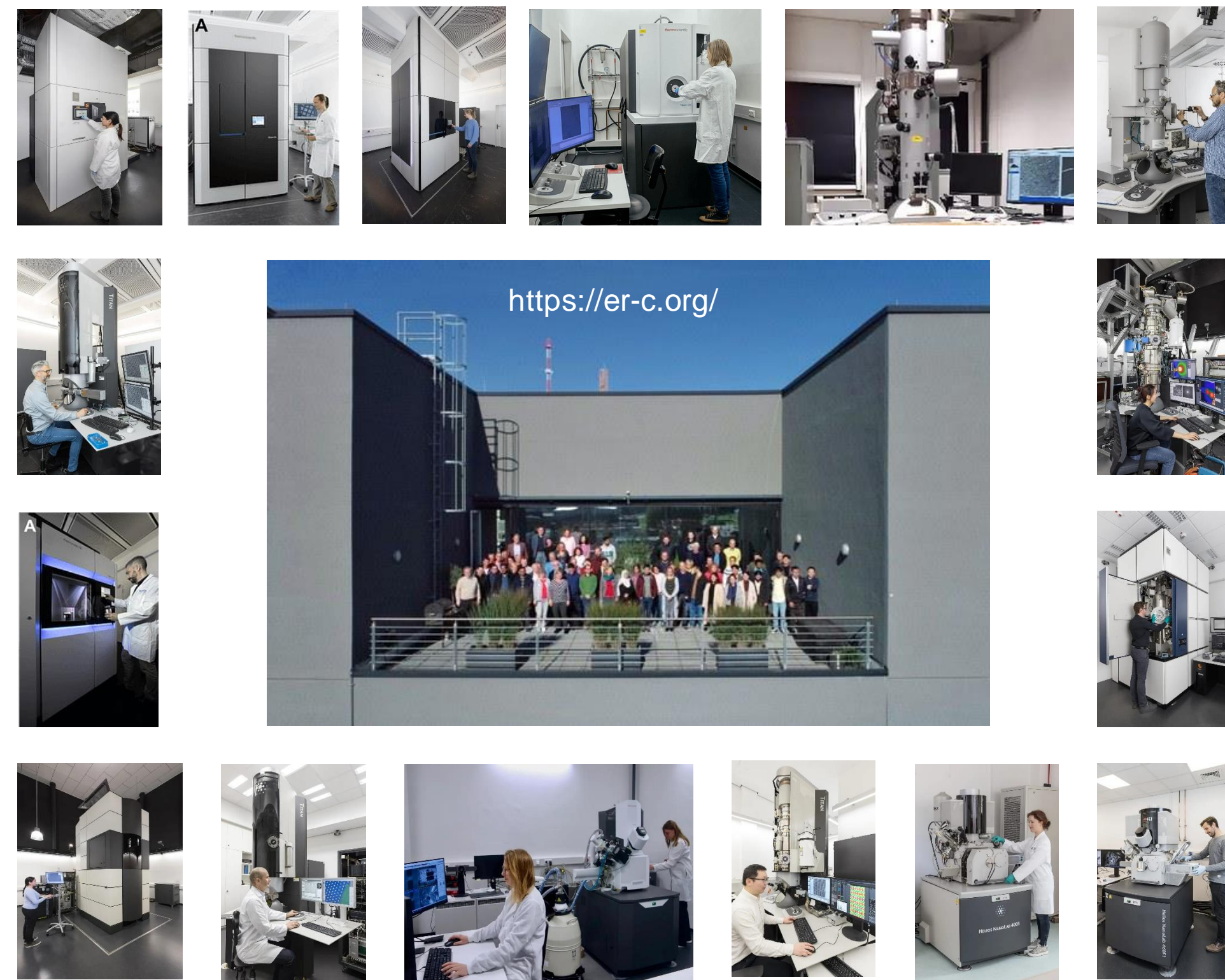
Ingest Data
Use our ingest GUI to store your session's data from the microscopes

Synchronize locally
Install and connect a native Nextcloud client

Large datasets
Transfer large folders or files

SampleDB
Learn how to access our electrons database via REST API

Report an issue
Contact us via our issue tracker



Requirements

- Manage heterogeneous data for heterogeneous user base
 - From basic to expert IT skills
 - Different tools for different people
 - Mix and match as desired
- Modular, extensible, versatile, simple, familiar
 - Integrate parts managed by different teams
- Can be large-scale – 100s of GB per dataset
- Performant data access
- Culturally acceptable: Users manage their own data in electron microscopy
 - Empower users
 - All day-to-day actions without admin intervention
- Integration with microscope systems
 - Often some older versions of Windows
 - Limit changes to setup, network

Acquisition

- Absorb the data coming from the detector and store it somewhere
- Inspect data quickly during session
- Keep notes, meta-information linked to raw data files
- Transfer to permanent storage location

Storage

- File system or object storage?
 - Most software at ER-C uses POSIX-like file system
- Maintain authentication and authorization throughout the system
- Prevent data loss
- Performant access

Access

- Collaboration, sharing, exchange
- Selective access control
- Internal, external
- Mobile devices
- Services
- External storage media
- Gateway for transfers
- Keep connection to metadata

Processing

- Fast data access for shared workstations
- Computation close to data
- Sync & Share for PCs, laptops

Publication

- Currently handled by users, external platforms
 - Zenodo
- Possible to set up public gateway

Archival

...future

SampleDB: Details with reference to file

Sample #250: Sample BLO file

Information

Action
Name: Sample BLO file
Material: BLO
Type: CTR
Comment: Sample preparation: David Hall, David Collier

Project
Name: libertem-test-data

Files
Title: https://er-c-data.fz-juelich.de/nextcloud/paths/nextcloud/nextcloud-test-data/default/...

Management: Network share for ingest

ER-C Data Management Self-Service

Manage ingest share for session "2022-03-22 my first session" in project "datamanagement"

Access using Windows Explorer
Map as network drive Z: on Windows

Access on Linux or OS X using a file manager

Windows network drive

Windows File Explorer showing network drives mapped to Z: and Y:.

Nextcloud web interface

Nextcloud web interface showing a list of files and folders.

Nextcloud sync client

Nextcloud sync client interface showing file synchronization status.

Search file references in SampleDB

SampleDB search results showing a table of objects with columns for Name, Date, User, Date 2, User 2, and Action.

Web-based browsing, link platform

ER-C Data Management Self-Service web interface showing project details and file browser.

Management web interface

ER-C Data Management Self-Service management interface showing project "libertem" details and sessions.

CIFS on shared systems

- Temporary accounts
- Reduce issues with cached credentials

CIFS

- Clients in fast local network

Browse contents

- Link to services
- Target for backlinks
- Access methods
- ELN integration
- File inspection

Nextcloud

- Web interface
- Sync & Share
- Mobile client

Nextcloud mobile client

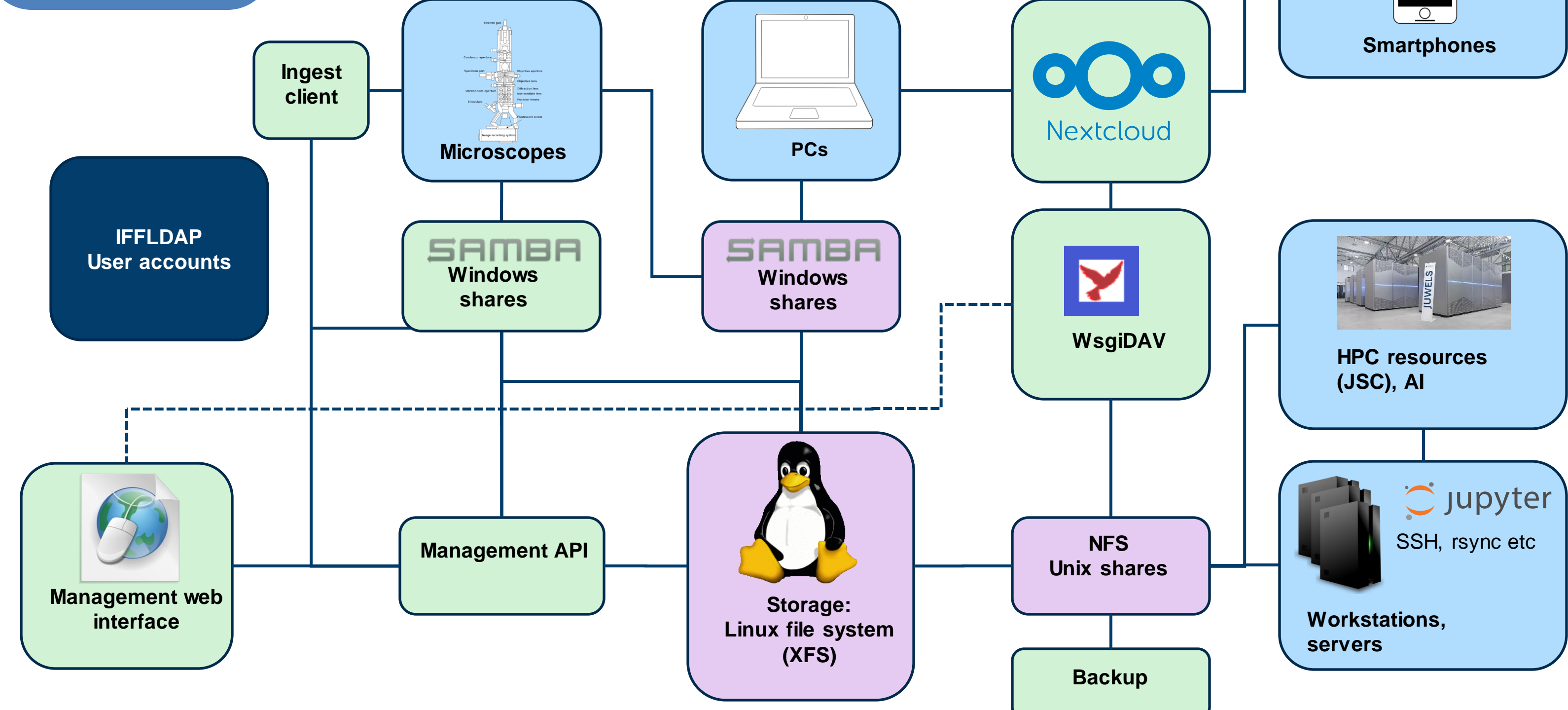
Nextcloud mobile client interface showing a list of files and folders.

JupyterHub file browser

JupyterHub file browser interface showing a list of files and folders.

LibeTEM web-based inspection

LibeTEM web-based inspection interface showing a microscopy image and analysis tools.



Interface for complex operations and elevated permissions

- Create projects
- Manage members
- Make data immutable

File system is source of truth for authorization

Groups based on ACLs

JupyterHub

Notebooks, plus launcher and proxy for web applications

File system access, including ACLs

```
(base) [weber@libf588 ~]$ nfs4_getfacl ~/er-c-data/adhoc/libertem
# file: /Users/weber/er-c-data/adhoc/libertem
# OWNER: rwaDTCcY
A::OWNER@:rwaDTCcY
A::20710:rwDTCcY
A::20831:rwaDTCcY
A::2601006:rwaDTCcY
A::GROUP@:rpxcY
A::EVERYONE@:rpxcY
A::FDI:OWNER@:rwaDTCcY
A::FDI:20710:rwDTCcY
A::FDI:20831:rwaDTCcY
A::FDI:2601006:rwaDTCcY
A::FDI:GROUP@:rpxcY
A::FDI:EVERYONE@:rpxcY
(base) [weber@libf588 ~]$ ls -ld ~/er-c-data/adhoc/libertem
drwxr-xr-x 5 weber libf 4096 Oct 23 09:13 /Users/weber/er-c-data/adhoc/libertem
(base) [weber@libf588 ~]$
```

Web-based data viewer

Web-based data viewer showing microscopy images and analysis tools.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823717 (ESTEEM3)

Cooperation within framework of **AIDAS**