iBridges: A comprehensive way of interfacing with iRODS

John Mc Farland, Christine Staiger, Tim van Daalen

iRODS User Group Meeting 2023
Chapel Hill North Carolina, USA

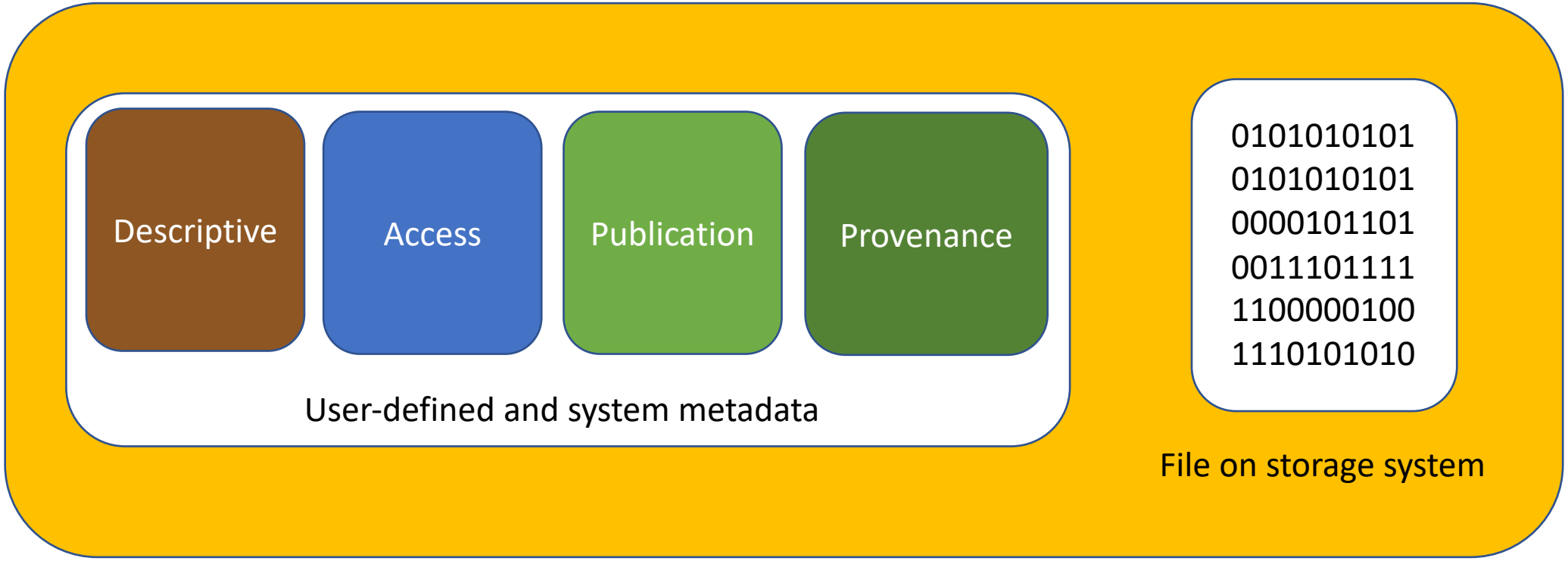
What would we need to master the data chaos?

- What if data objects
  - Could tell us where their copies are stored?
  - Could tell us which other data is derived from them?
  - Would carry metadata describing how they were generated?
  - Could tell us in which state they are (version, published, volatile)?

- Other metadata than system metadata → user-defined metadata
- System and user-defined metadata that cannot be separated from the data object
- A service that shows us the data object and not only the files

iRODS provides Data objects

Data object = File + metadata + identifier



Keeping data and metadata tied together: iRODS

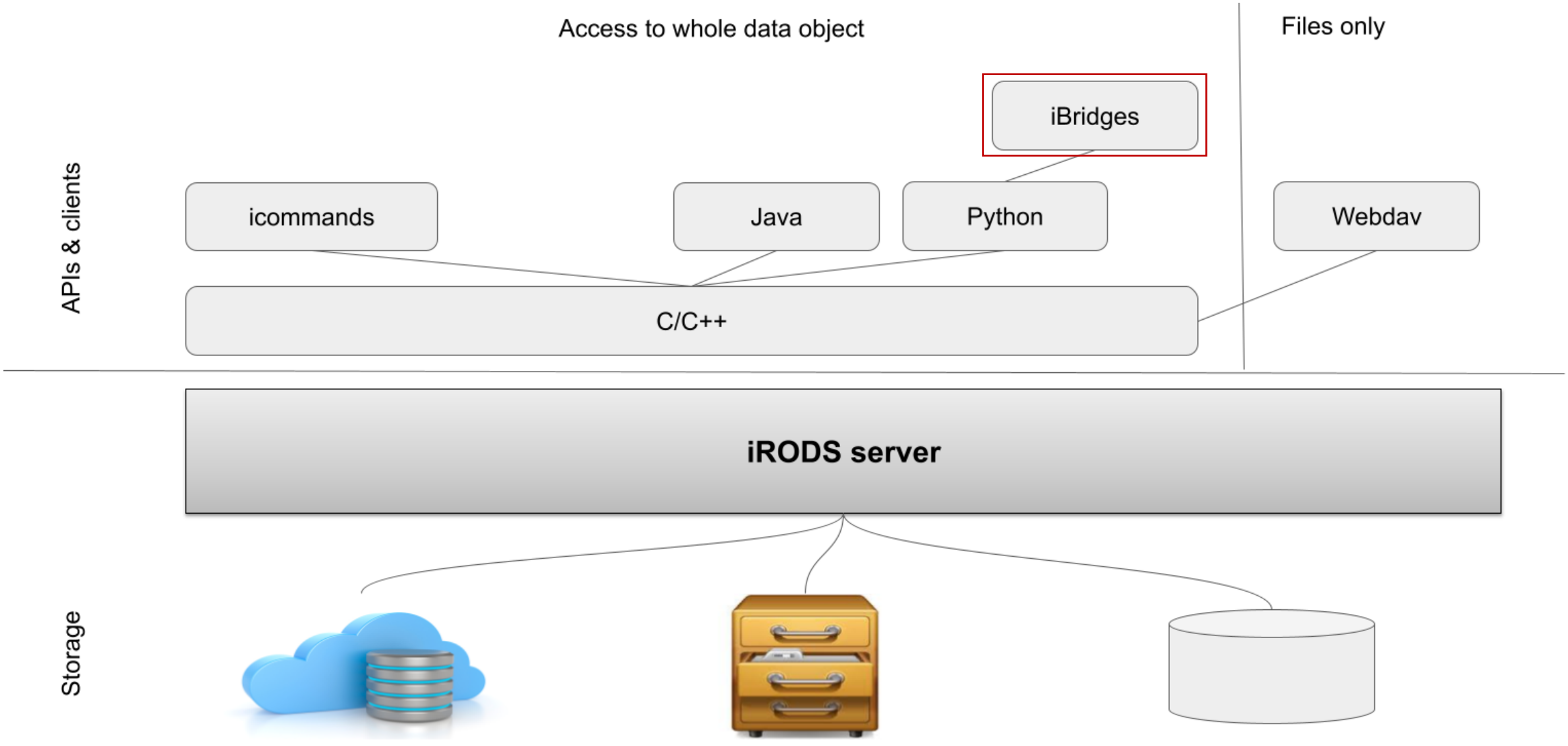
iRODS is great, but …

- We need an easy interface for researchers to interact with their data

  → GUI

- Use case requirements
  - NPEC project, Wageningen UR:
    - standalone software, needs to be able to run on computers provided by suppliers for laboratory equipment → no influence on preinstalled software
    - Large data uploads; ~500GB per upload
    - Continuous data uploads (synchronisation from local source to iRODS)
- GUI
  - Supported on major platforms: Windows, macOS and Linux
  - Respects and does not to interfere with any possible policies on the iRODS servers
  - Easily interface with other data services on the client side

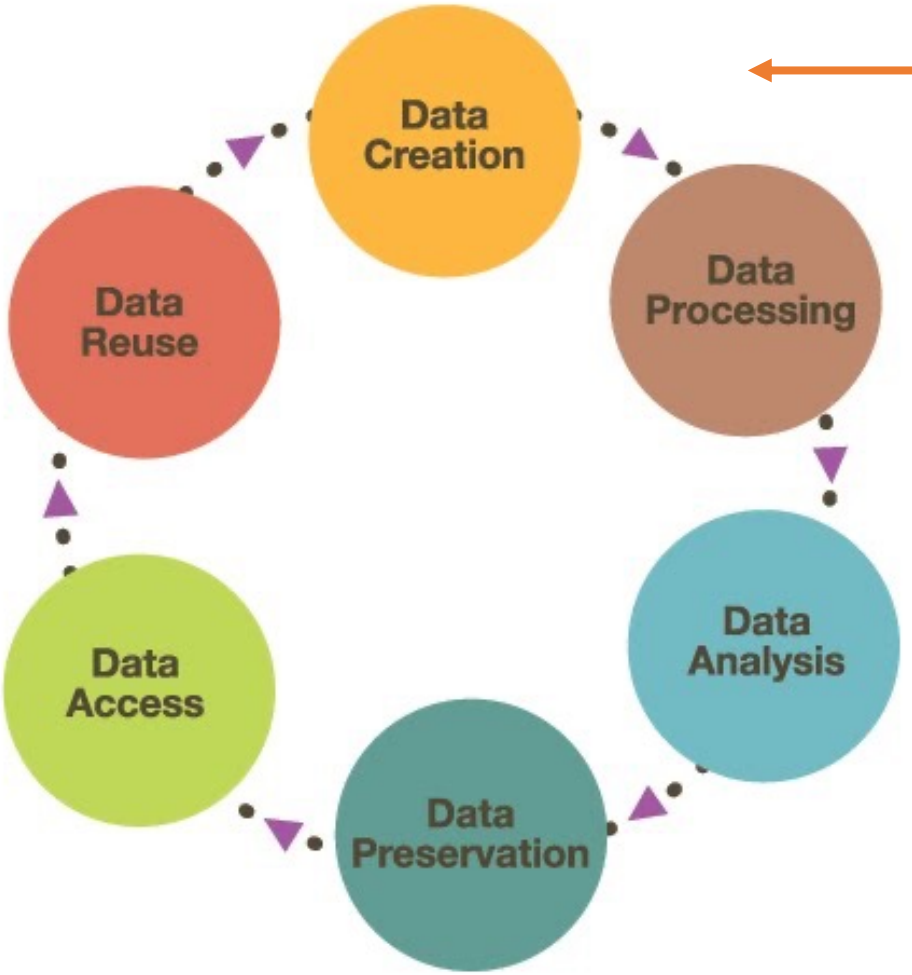
YODA: Accessing data objects



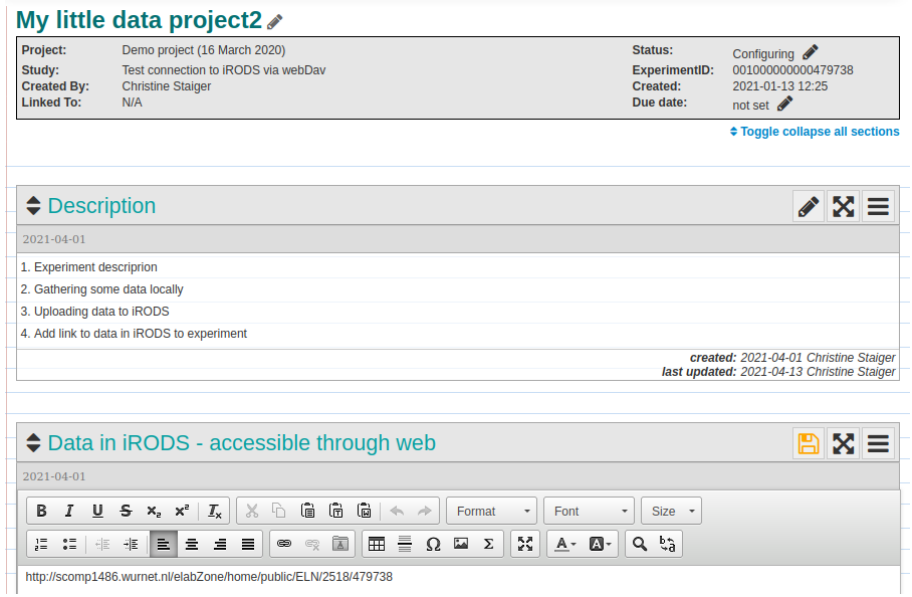
Interfacing with other services

- Client-side integration
  - Light-weight integration
  - Code should potentially also be easy to understand to enable scientific programmers to provide their group with a different view on their data
  - Integrate iRODS and other services on API level
- Examples
  - Integration of iRODS with an electronical labnotebook server
  - Integration of iRODS with an audio transcription service

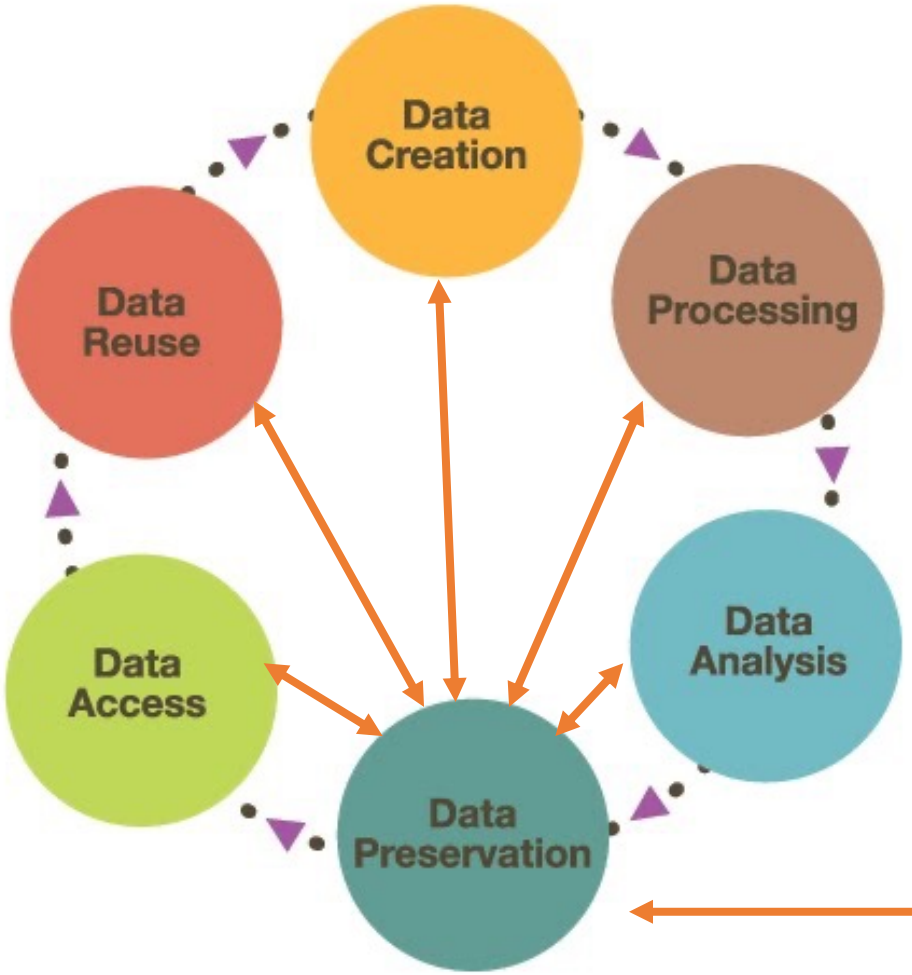
From data creation to archiving: ELN



ElabJournal: Describe how data is generated in a lab (Experiment)



From data creation to archiving: Yoda



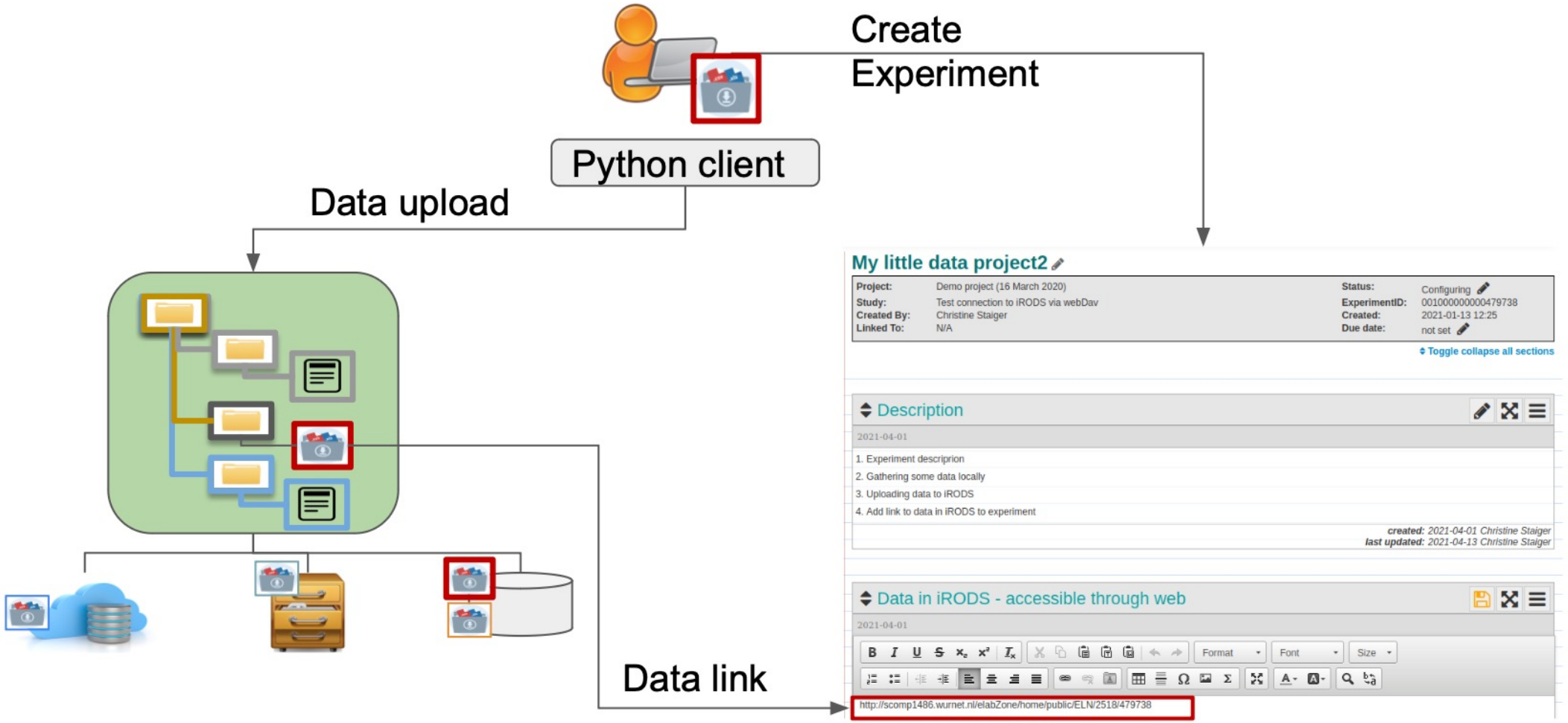
- Based on iRODS (Talk by Lazlo Westerhoff in the afternoon)
- Provides data steward workflow to
  - Ensure quality of archived data, data curation
  - Transfer responsibility from researcher to data manager

- No matter in which state your data is, Yoda is the route and guide to safe archiving

- Check out archived data to any stage in which the project is

Yoda: Preparing data for archiving and properly archive data

The workflow



Demo time

Future work & the Team

- Unit tests for backend (in progress)
- Data synchronisation (in progress)
- Improving performance of data transfers through the python API (in progress)
- Code optimalisation of the front-end
- Packaging software as executable
- Plugin structure for interfacing with other data services
- Extend iRODS ticket functionality

Contribute or test:
https://github.com/chStaiger/iBridges-Gui

The team and special thanks to:

University of Groningen:
John Mc Farland, Simona Stoica

Utrecht University:
Maarten Schermer, Christine

Wageningen University and Research:
Jos Tersteeg, Tim van Daalen