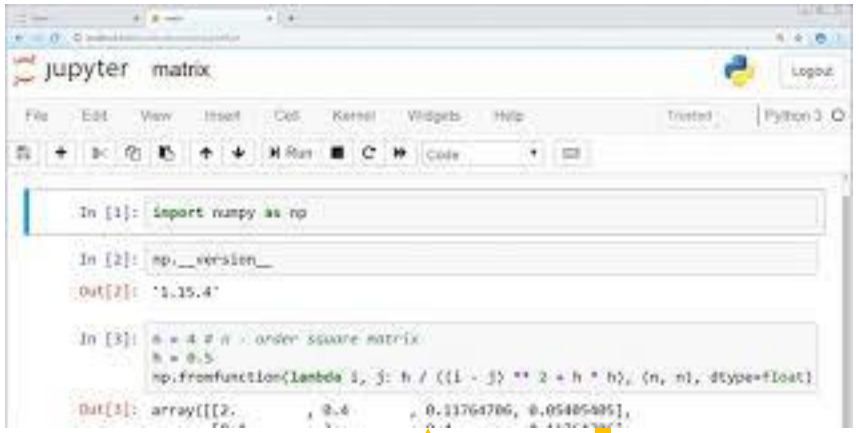
iBridges

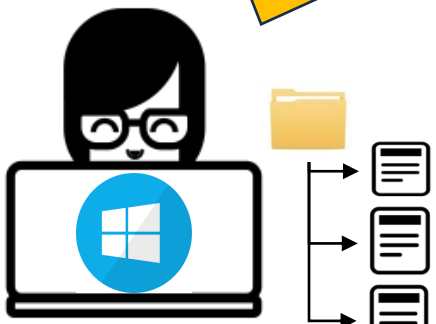
How easy can we make iRODS for researchers?

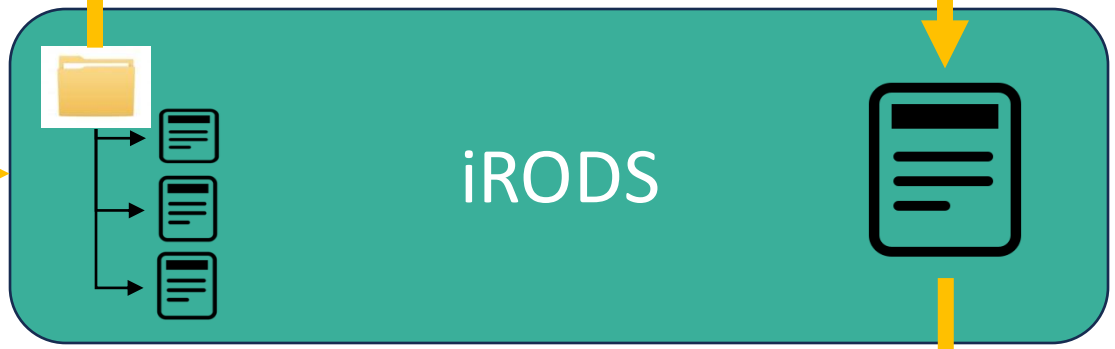
Raoul Schram, Maarten Schermer, Christine Staiger

Collaboration on Data

Can you **run your analysis again on the new data** and **give me the results** for my presentation?







What is iBridges?

- Python API
- Commandline Interface + shell with auto-completion
- Graphical User Interface

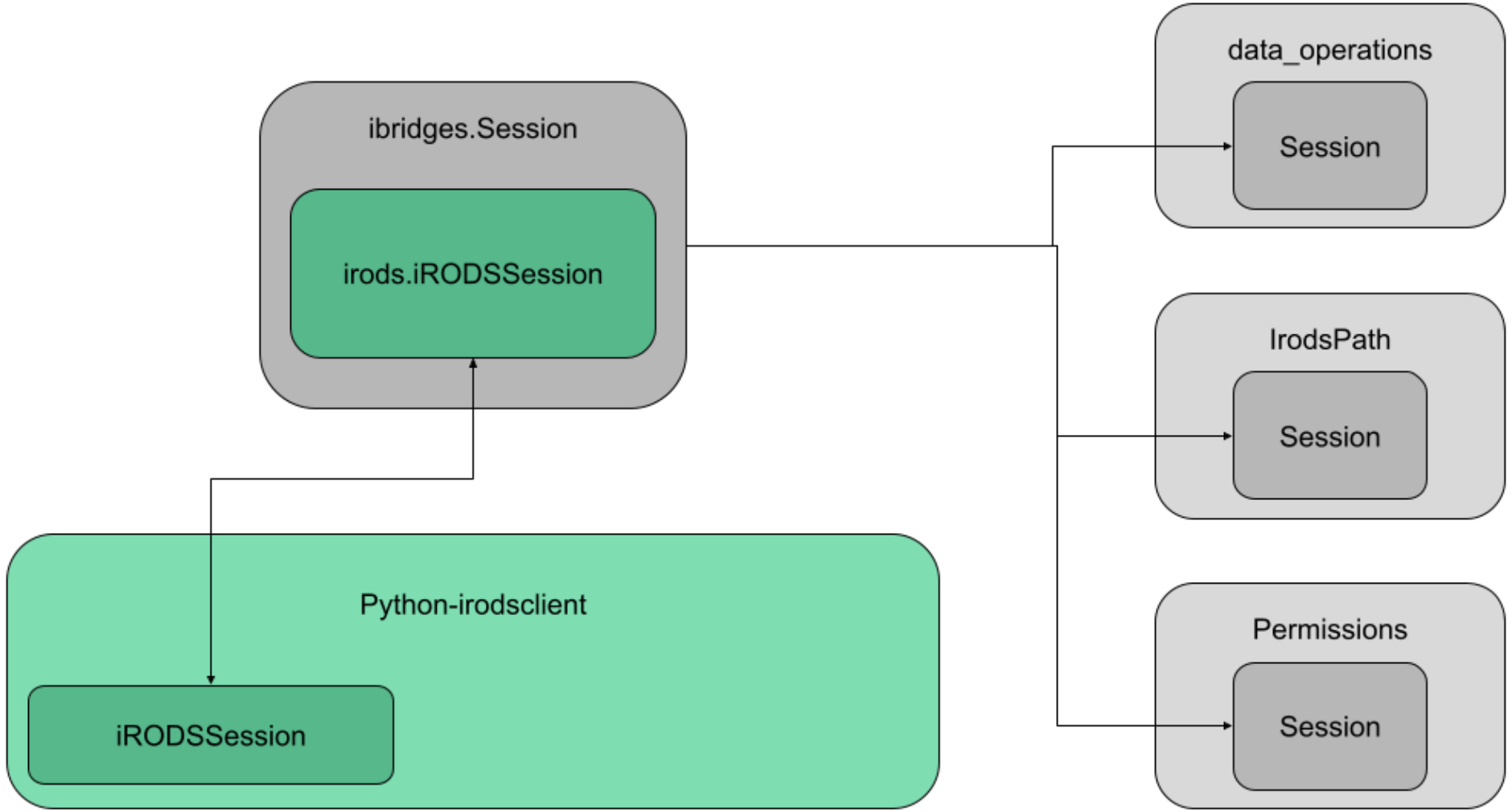
Meant for:

- Researchers of all backgrounds
  - Different technical expertises
  - Windows, MacOs and linux
  - Managing files and their metadata

Use cases

- I am a **researcher** and I want to **interactively explore my data** in iRODS
- I am a **researcher** and I need to pull in a **subset of my data** into my **compute workflow**
- I am a **lab technician**, my equipment produces a lot of data which I need to **upload to iRODS from Windows** machines
- I am a **data manager** and I need to **monitor the inflow of data** from other people or equipment

The iBridges API and the PRC



Highlights from the last year

- iBridges shell: tab-completion for iRODS paths, local paths and commands

- iBridges API:
  - IrodsPath class
  - Metadata
    - Access from IrodsPath directly
    - Setters for key, value and units
  - Include metadata in data transfers
  - Stream data objects by their IrodsPath
  - Improved search
  - Introduced `cwd`

Highlights from the last year

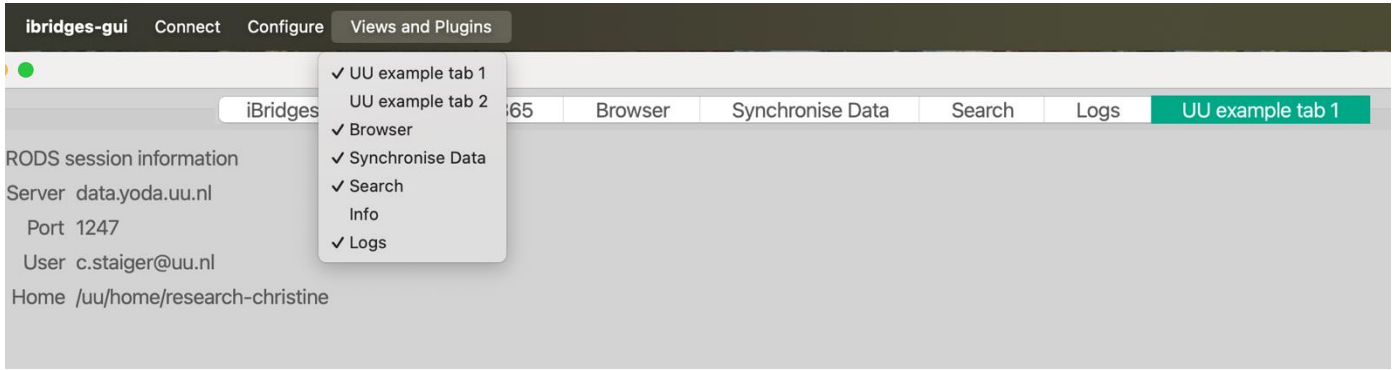
- iBridges shell: tab-completion for iRODS paths, local paths and commands

- iBridges API:
  - IrodsPath class
  - Metadata
    - Access from IrodsPath directly
    - Setters for key, value and units
  - Include metadata in data transfers
  - Stream data objects by their IrodsPath
  - Improved search
  - Introduced `cwd`

- iBridges CLI:
  - Aliases for different iRODS servers
- iBridges GUI:
  - PyQt → PySide
  - Executables and scripts to build them
  - New entrypoint: `ibridges gui`
  - Include aliases from CLI
- Lots of documentation, examples and tutorials!

For developers: Plugins

- iRODS server templates
- Customised views to the GUI
- New commands to the CLI and shell

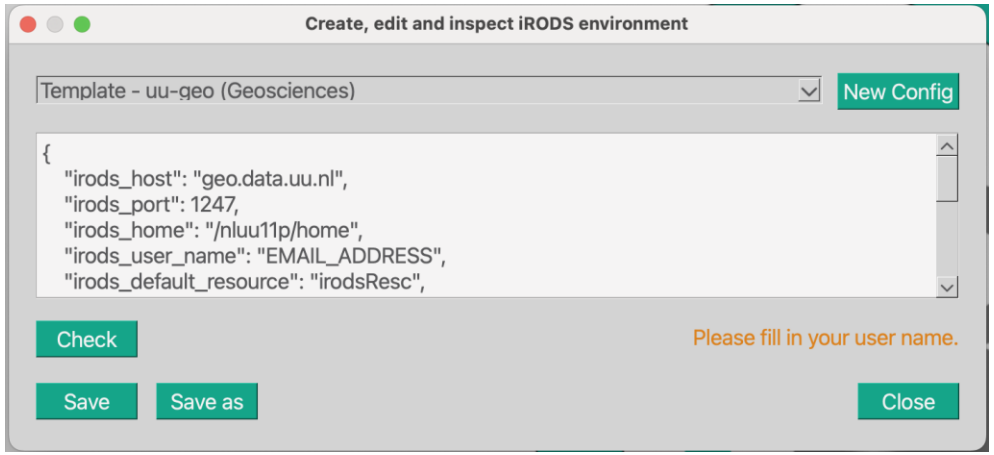


- Example plugins with small set of instructions
  - GUI: https://tinyurl.com/ibridges-gui
  - CLI/Shell: https://tinyurl.com/ibridges-cli
  - Configurations: https://tinyurl.com/ibridges-configs

```
ibshell:research-christine> help

Documented commands (type help <topic>):
========================================
EOF  download  info  meta-add  meta-list  pwd   rm      shell  tree
cd   help      ls    meta-del  mkcoll     quit  search  sync   upload

Undocumented commands:
======================
del  l  list  remove

ibshell:research-christine> info
Host      : data.yoda.uu.nl
Port      : 1247
Version   : 4.2.12
```

Challenge 1: First time login

- Plugin for iRODS server templates
- Guidance for users what to fill in and how to store it
- GUI: Check if file contains server-compatible parameters




```
staig001@busoni ~ % ibridges setup --list
Utrecht University templates
----------------------------


uu-youth     - YOUth Cohort Study
uu-geo       - Geosciences
uu-i-lab     - Humanities, Law, Economics, Governance, Open Societies
uu-dgk       - Veterinary Medicine, Medicine
uu-science   - Science
uu-fsw       - Social and Behavioral Sciences
uu-its       - University Corporate Offices
uu-surf      - Yoda instance hosted at SURF
staig001@busoni ~ % ibridges setup uu-geo
email_address: my_mail@uu.nl
```

Challenge 2: Several iRODS configurations

- Aliasing and matching with cached password

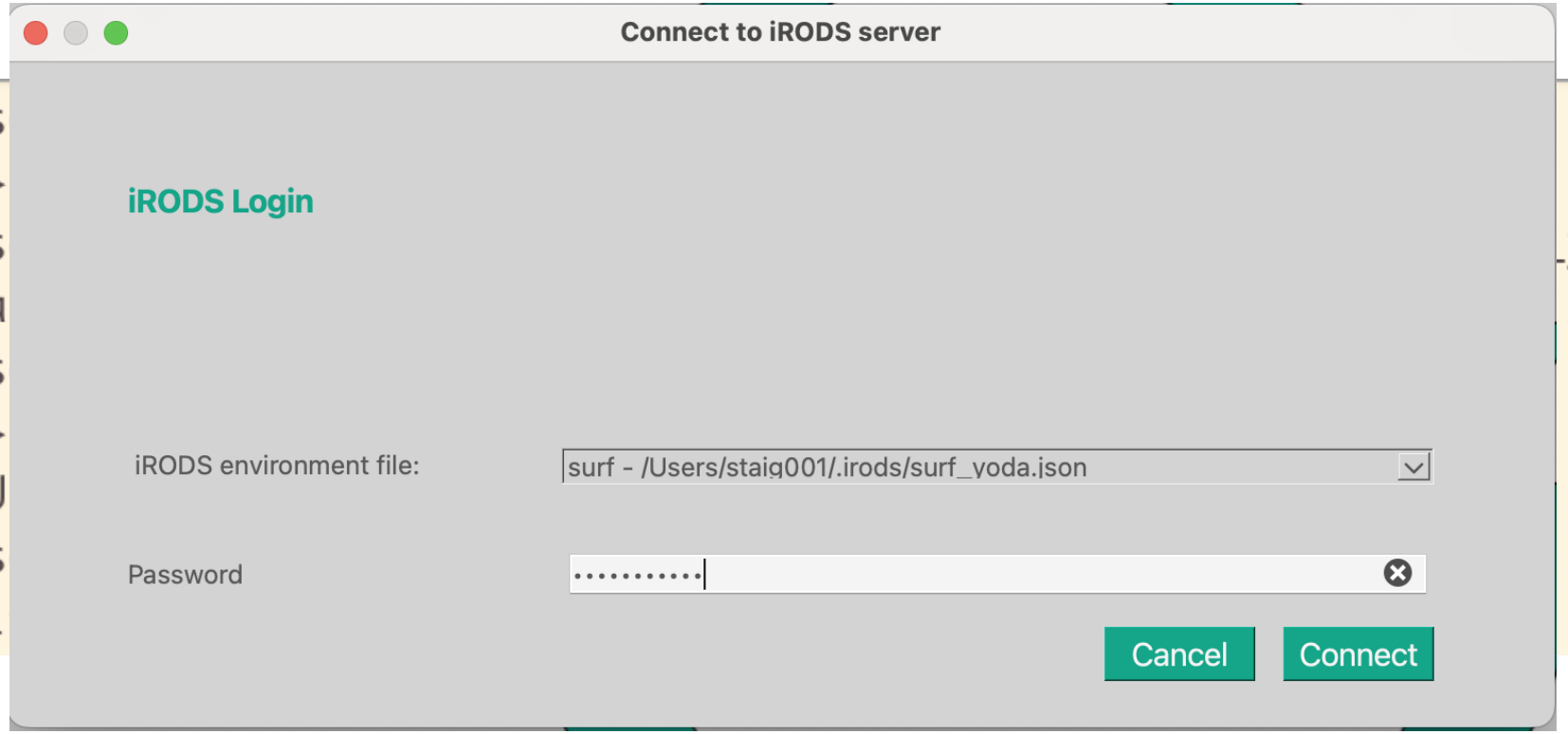
```
staig001@busoni ~ % ibridges alias
* default -> /Users/staig001/.irods/irods_environment.json
staig001@busoni ~ % ibridges alias surf /Users/staig001/.irods/surf_yoda.json
Created alias 'surf'
staig001@busoni ~ % ibridges alias
* default -> /Users/staig001/.irods/irods_environment.json
  surf -> /Users/staig001/.irods/surf_yoda.json
staig001@busoni ~ % ibridges init surf
ibridges init was succesful.
```

Challenge 2: Several iRODS configurations

- Aliasing and matching with cached password

```
staig001@bus
* default ->
staig001@bus
Created alia
staig001@bus
* default ->
  surf -> /U
staig001@bus
ibridges ini
```

```
_yoda.json
```



Challenge 3: iRODS Paths

- IrodsPath and CachedIrodsPath
- Pathlib like functionality: *joinpath, parent, parts, relative_to*
- Shortcuts: `~` and `.`
- Information: size, checksum, metadata
- Streaming

Challenge 3: iRODS Paths

```
In [1]: from ibridges.interactive import interactive_auth

In [2]: from ibridges import IrodsPath

In [3]: session = interactive_auth(irods_env_path="/Users/staig001/.irods/surf_yoda.json")

In [4]: coll_path = IrodsPath(session, "my_books")

In [5]: coll_path.size
Out[5]: 6444527

In [6]: obj_path = coll_path / "DonQuixote.txt"

In [7]: print(obj_path.meta)
 - (key: 'author', value: 'Miguel de Cervantes', units: '')
```

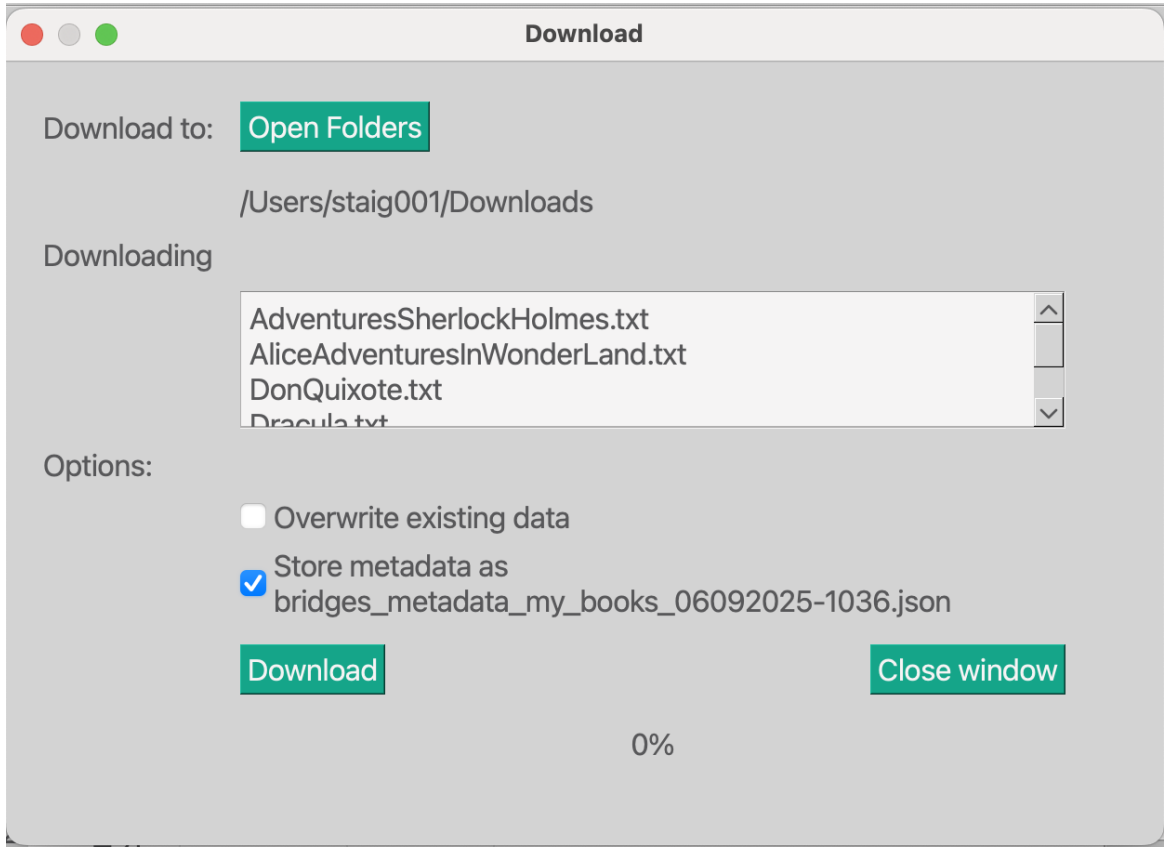
Challenge 4: Metadata manipulation

- Editing keys, values and units
- Find metadata item(s)
- Access metadata by key

```
In [4]: coll_path = IrodsPath(session, "my_books")

In [5]: print(coll_path.meta)
 - (key: 'contains', value: 'text files', units: '')
 - (key: 'data_type', value: 'collection', units: '')
 - (key: 'data_type', value: 'demo_data', units: '')

In [6]: coll_path.meta.find_all("data_type")
Out[6]:
[<MetaDataItem (data_type, demo_data, )>,
 <MetaDataItem (data_type, collection, )>]

In [7]: meta_item = coll_path.meta["contains"]

In [8]: meta_item.value = "text_files"

In [9]: print(coll_path.meta)
 - (key: 'contains', value: 'text_files', units: '')
 - (key: 'data_type', value: 'collection', units: '')
 - (key: 'data_type', value: 'demo_data', units: '')
```

Challenge 5:
Up and downlaod including metadata

- GUI: only downlaod
- CLI:
  - Upload
  - Download
  - Synchronisation



Challenge 5:
Up and downlaod including metadata

```
ibshell:research-christine> sync Downloads/my_books irods:my_books_backup --metadata Downloads/my_books_meta.json
100%|                                                                                | 6.15M/6.15M [00:14<00:00, 430kB/s]
ibshell:research-christine>
ibshell:research-christine> list -m my_books_backup
/uu/home/research-christine/my_books_backup:
 - (key: 'contains', value: 'text_files', units: '')
 - (key: 'data_type', value: 'collection', units: '')
 - (key: 'data_type', value: 'demo_data', units: '')

  /uu/home/research-christine/my_books_backup/AdventuresSherlockHolmes.txt
 - (key: 'author', value: 'Conan Doyle', units: '')
```

Challenge 5: Metadata Up and Download

- We tried to minimise the “nestedness”

- Entries for
  - Relative path to the selected iRODS root
  - Checksum
  - Name
  - iRODS ID
  - Metadata as list of key-value-units triples

Challenge 5:
Metadata Up and Download

- JSON
  - ibridges_metadata_version : "1.0"
  - recursive : true
  - root_path : "/uu/home/research-christine/my_books"
  - items
    - 0
      - rel_path : "."
      - type : "collection"
      - name : "my_books"
      - irods_id : 195115
      - metadata
    - 1
      - rel_path : "AdventuresSherlockHolmes.txt"
      - type : "data object"
      - name : "AdventuresSherlockHolmes.txt"
      - irods_id : 195120
      - checksum : "sha2:Y8VmrN3ikAZyEXOR/u+4XtbOL8R386UkE49HxUlkd7k="
      - metadata

Challenge 6: Searching for data

Search in /uu/home/research-christine

Search Wildcard is: %

Obj/Coll name %book%/%

Checksum

Search by Metadata

Key	Value	Units
author		

Case sensitive

Search | Clear Results | Select all | Download Selection

	Type	Path	Size [bytes]	Created	Modified
4	-d	/uu/home/research-christine/my_books/DonQuixote.txt	2284484	10-03-2025	10-03-2025
5	-d	/uu/home/research-christine/my_books/Frankenstein.txt	426421	10-03-2025	10-03-2025
6	-d	/uu/home/research-christine/my_books_backup/AliceAdventuresInWonderLand.txt	148273	09-06-2025	09-06-2025
7	-d	/uu/home/research-christine/my_books/AdventuresSherlockHolmes.txt	594238	10-03-2025	10-03-2025
8	-d	/uu/home/research-christine/my_books/Phantasmagoria.txt	109546	10-03-2025	10-03-2025
9	-d	/uu/home/research-christine/my_books_backup/AdventuresSherlockHolmes.txt	594238	09-06-2025	09-06-2025

Distinction between Objects and Collections

- Separate tables for data objects and collections → all native APIs have two modes for commands

OBJ ID	OBJ NAME	COLL NAME	META_DATA	
1234234	Data.txt	/zone/home…	(KEY, VALUE, UNITS)	

COLL ID	COLL NAME		META_COLL	
1234234	/zone/home…		(KEY, VALUE, UNITS)	

?

- collection path → Collections
Filter data queries on:
- collection path → Data objects
- data name → Data Objects

Our Users

- Researchers at Utrecht University
- Netherlands Plant Eco-phenotyping Centre, https://www.npec.nl/
  - Ingest of large data through the GUI
- UNLOCK, https://m-unlock.com/
  - User manuals to work with data
- EXPANSE, https://expanseproject.eu/
  - Giving access to data to external and international project partners

Pypi suggests there must be more users and we would like to get to know you!

Join us on November 27th for the first iBridges user meeting!!!

Many thanks to

Tim van Daalen, Jasper Koehorst

Wageningen University and Research, NL

John Mc Farland, Simona Stoica, Jelte Nimoth

Rijksuniversiteit Groningen, NL

Terrell Russel

iRODS

Mark Schenk

Delft University of Technology, NL

Jörg Steinkamp

Johannes Gutenberg-University Mainz, DE

Zimbo Boudewijns, Adam El Kassimi, Dawa Ometto,

Maarten Schermer, Raoul Schram, Sietse Snel

Utrecht University, NL

Join us!

iBridges User Meeting November 27th

And on Github

https://github.com/iBridges-for-iRODS

edu.nl/yax8p