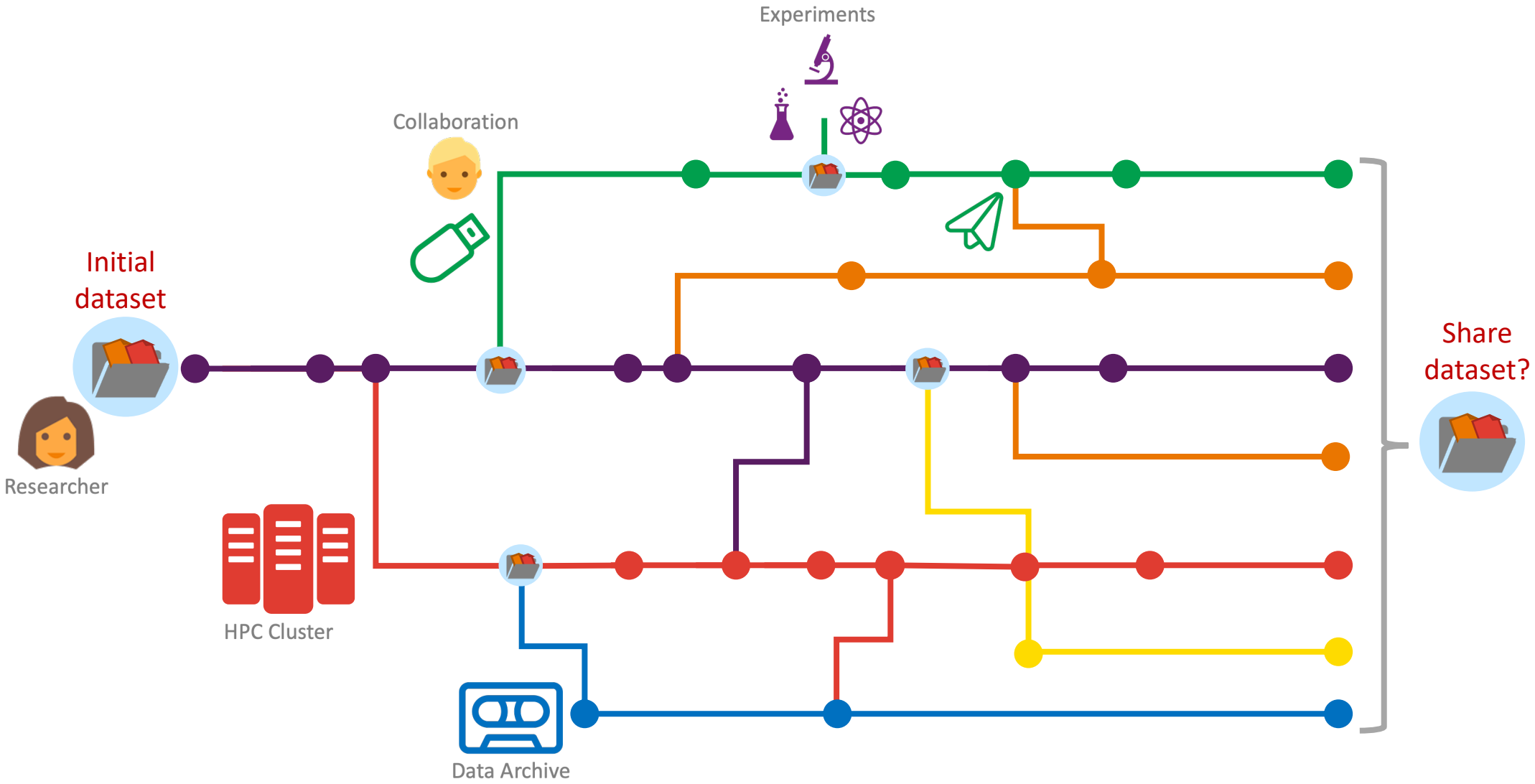
Introduction RDM with iRODS and HPC

Start at 9:00

Outline

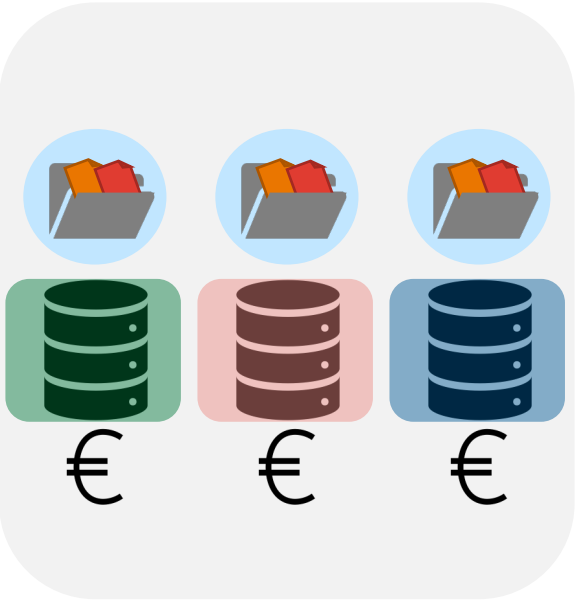
- The problem with data
- iRODS can help

Data, what's the problem?



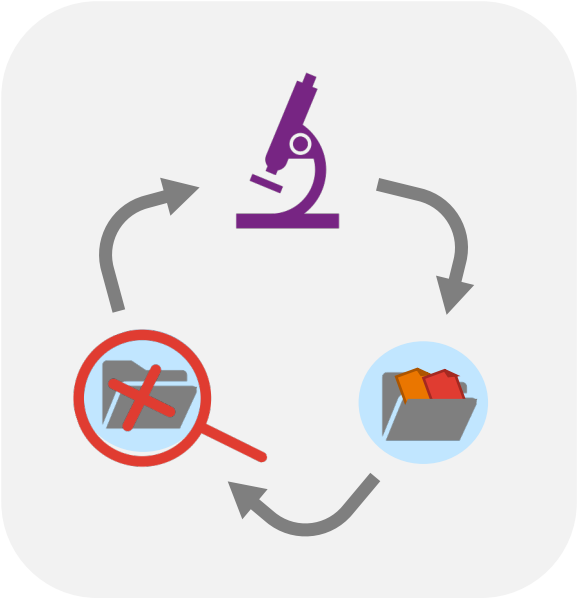

Share
dataset?

Consequences of bad research data management

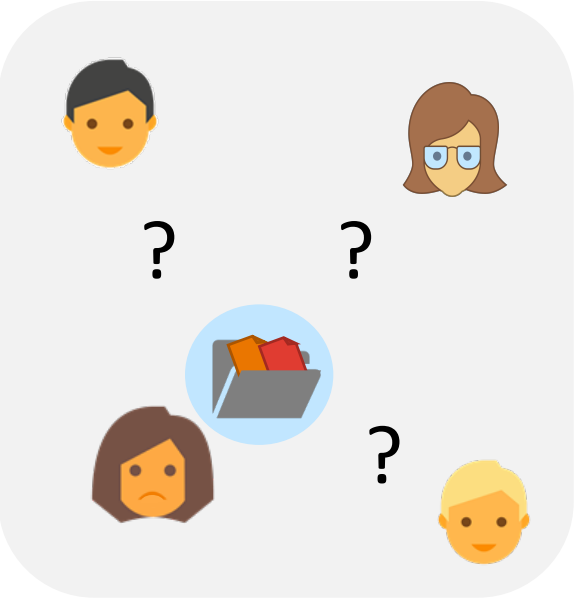


No cost effective data storage

Data gets lost by disaster or loss of context

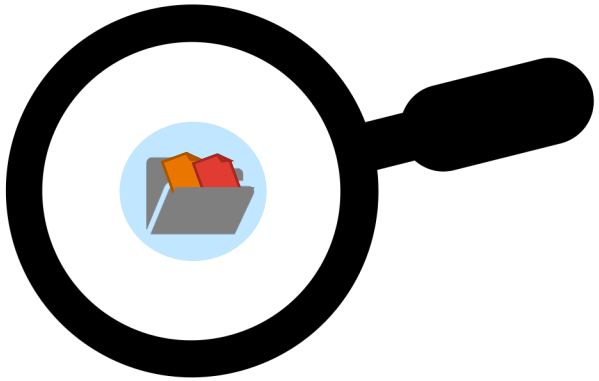


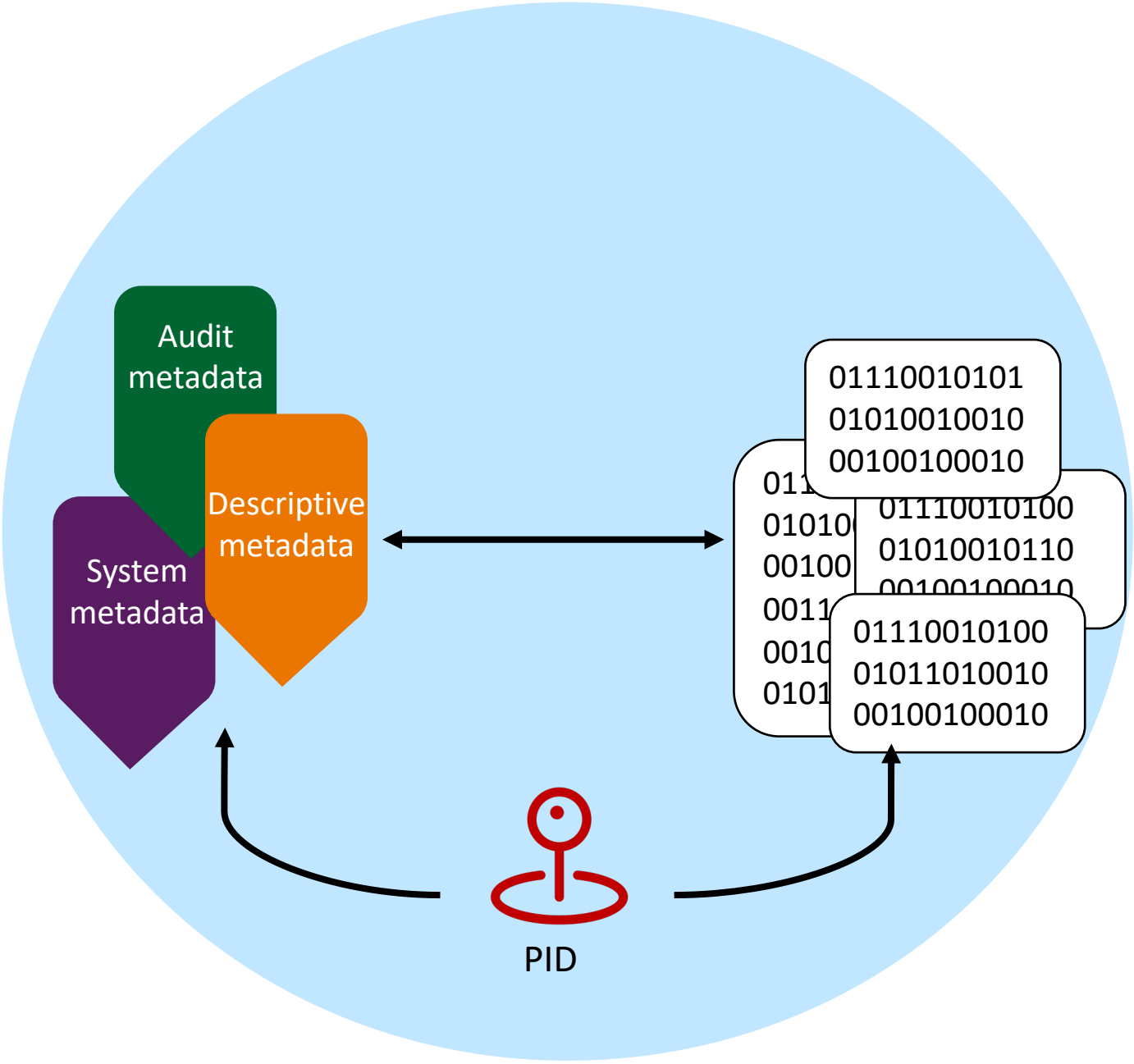
Redoing experiments and no interoperability



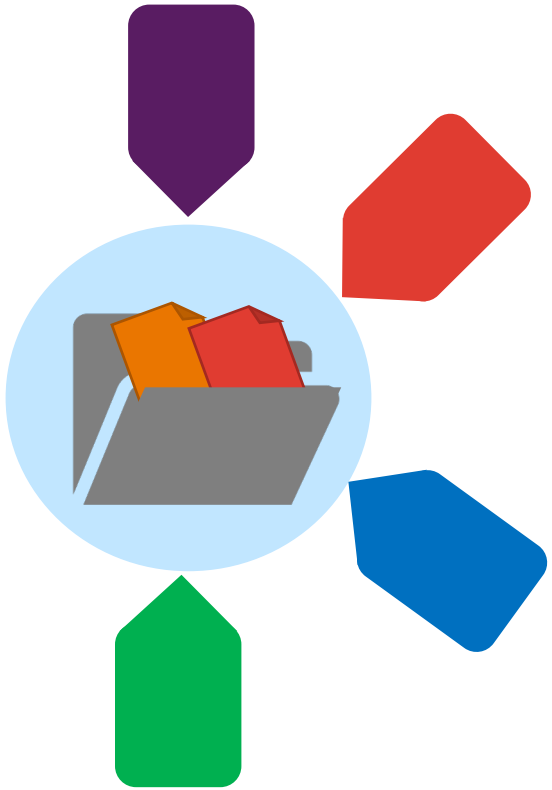
Not able or fear to share / publish data

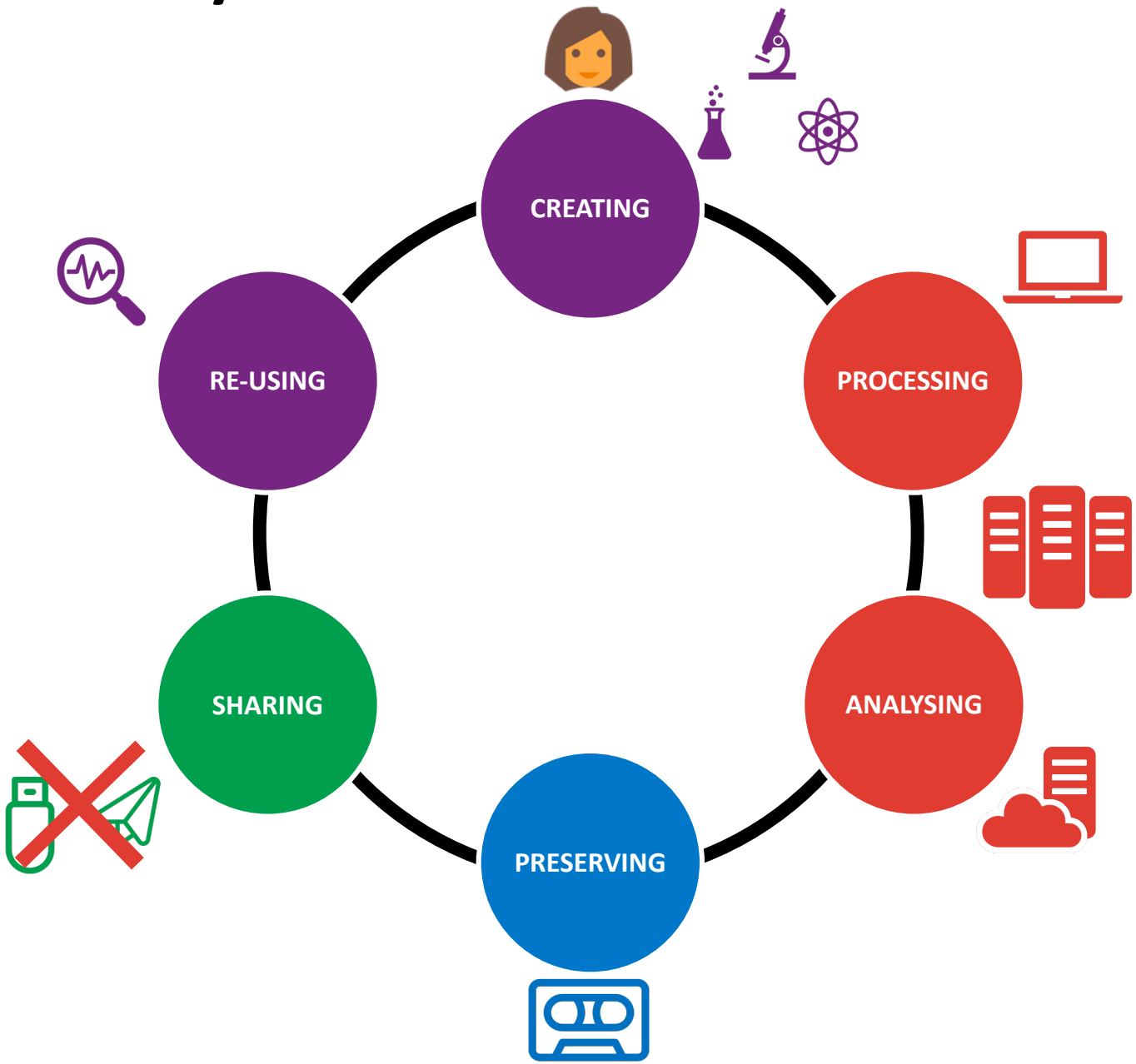
Data, what is it?



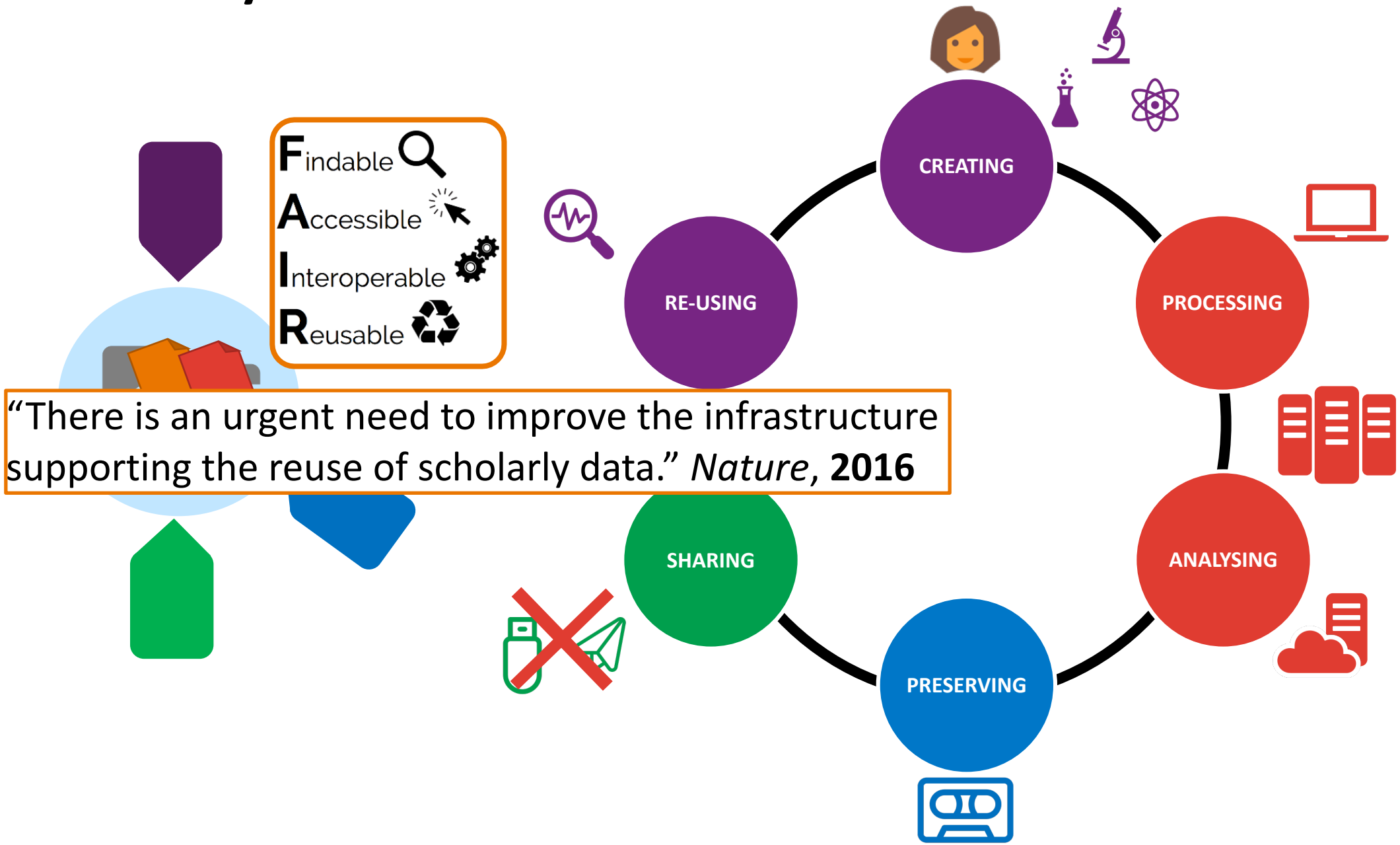


Data Life Cycle is described by metadata





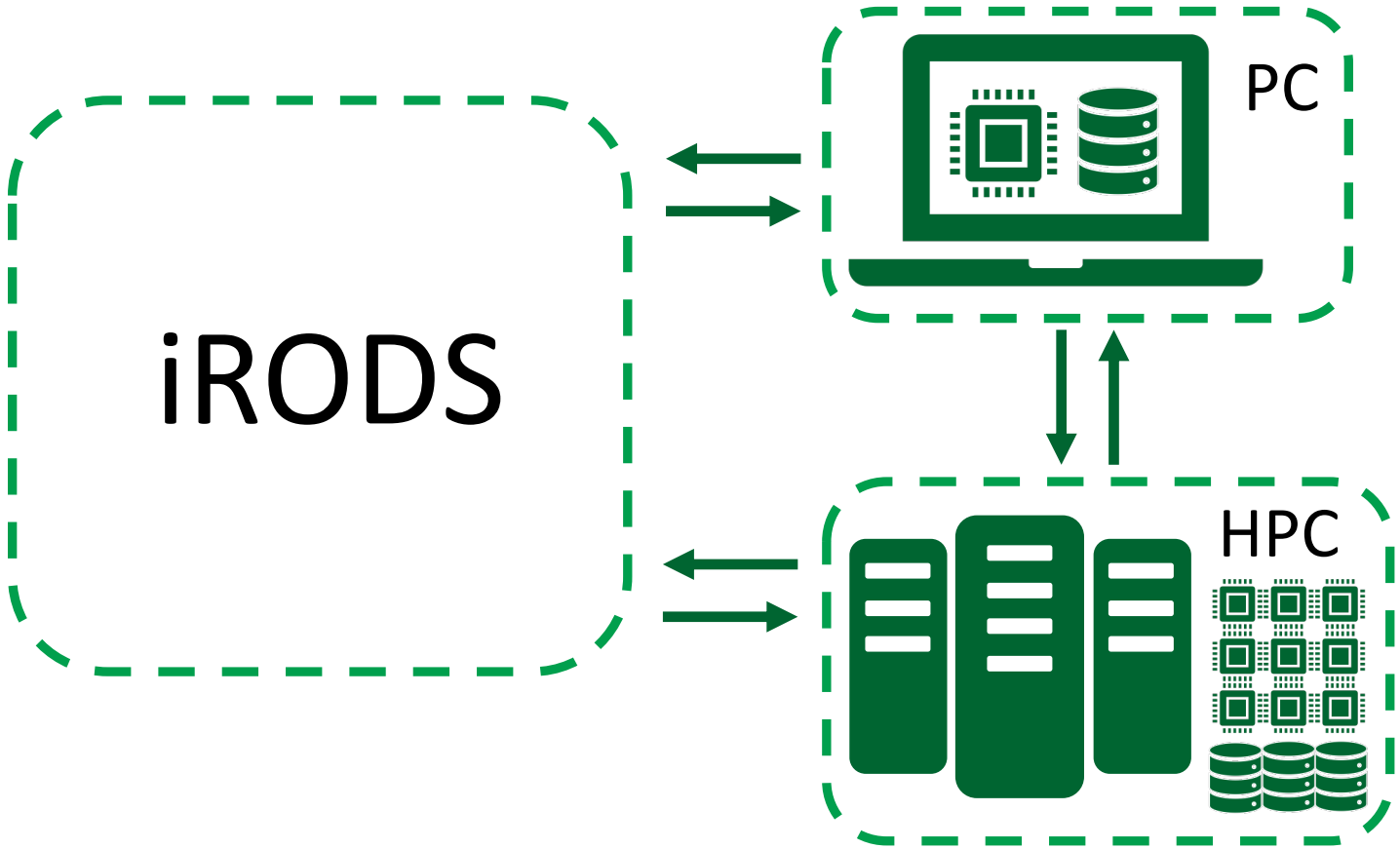
Data Life Cycle is FAIR due to metadata



"There is an urgent need to improve the infrastructure supporting the reuse of scholarly data." *Nature*, **2016**

Data Life Cycle can be fully supported by iRODS

Data storage from PC to HPC



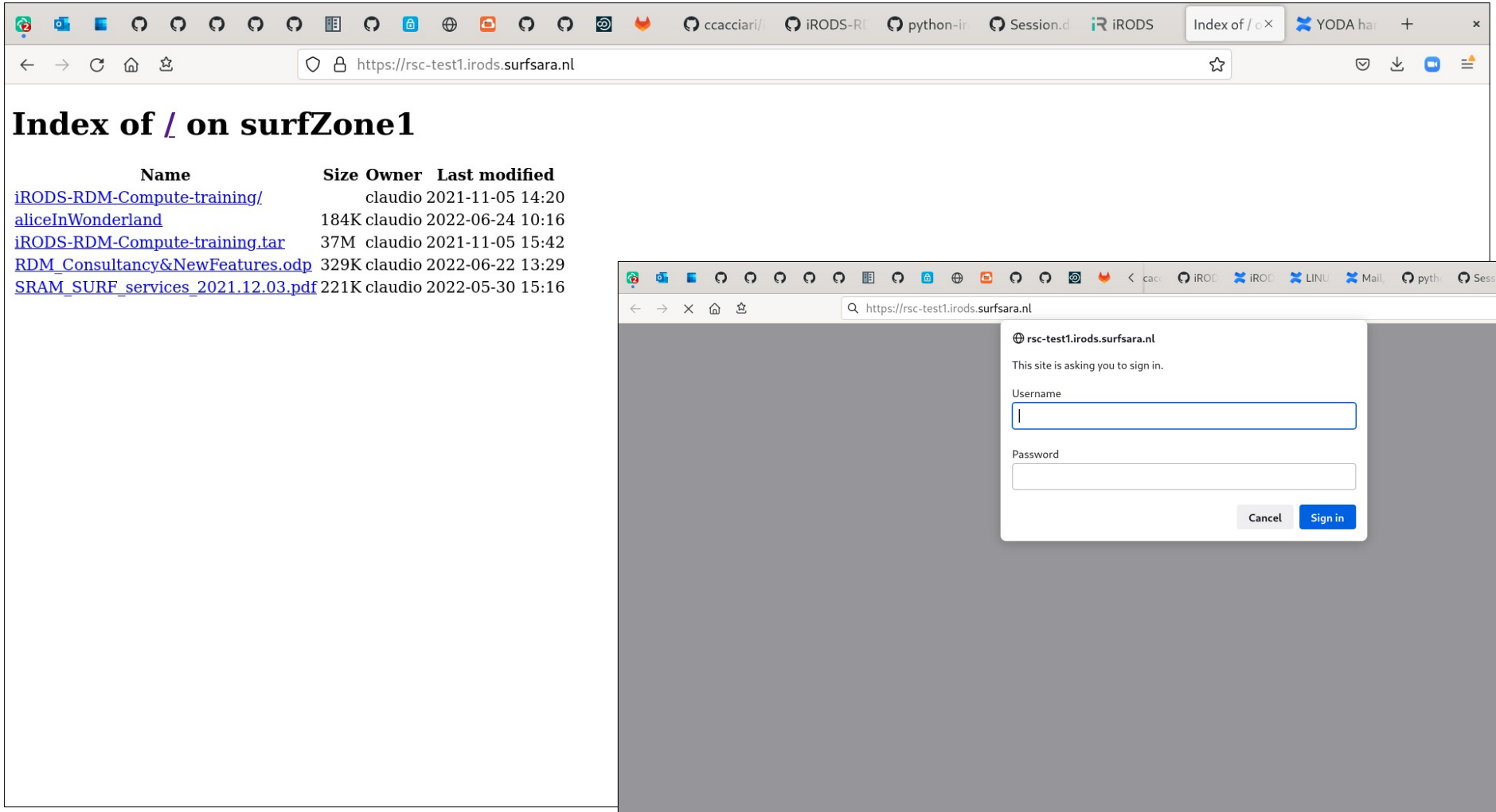
iRODS clients

- iRODS offers multiple clients
  - https://irods.org/clients
- Many are developed by the community and their level of maturity varies a lot.
- We will focus on the two most reliable and supported:
  - the command line interface: icommands,
  - the python client API
- And on the WebDAV interface
  - WebDAV is a standard protocol and there are clients for all the main operating systems: Windows, Linux, Mac.

Hands on for data processing with iRODS

- Learn how to handle data with iRODS using a webdav client
- Learn how to handle data via iRODS python api

WebDAV


https://rsc-test1.irods.surfsara.nl
Index of / on surfZone1
Name Size Owner Last modified
iRODS-RDM-Compute-training/ claudio 2021-11-05 14:20
aliceInWonderland 184K claudio 2022-06-24 10:16
iRODS-RDM-Compute-training.tar 37M claudio 2021-11-05 15:42
RDM_Consultancy&NewFeatures.odp 329K claudio 2022-06-22 13:29
SRAM_SURF_services_2021.12.03.pdf 221K claudio 2022-05-30 15:16

https://rsc-test1.irods.surfsara.nl
rsc-test1.irods.surfsara.nl
This site is asking you to sign in.
Username
Password
Cancel
Sign in
rsc-test1.irods.surfsara.nl

iRODS Python client API

- https://github.com/irods/python-irodsclient
- Version 1.1.3
- Tested against iRODS 4.2.11
- Why an API? Why not a CLI or GUI client?
  - There are some python clients, based on this API, developed by the community
    - PROS: a CLI client is easier to use
    - CONS: a CLI client is harder to maintain