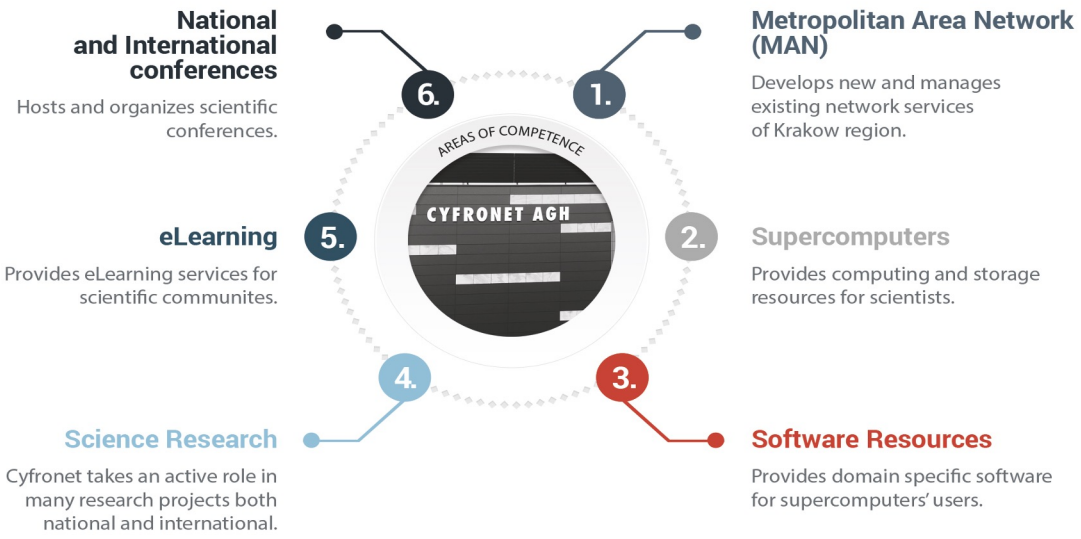
Academic Computer Centre
CYFRONET AGH

Roksana Wilk, ACC Cyfronet AGH
Head of Data Processing Laboratory, Cyfronet's EOSC and Open Science affairs rep

Skills and competences in Open Science

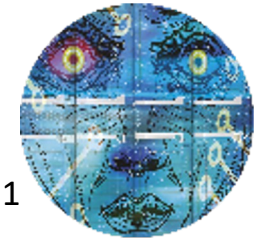
Academic Computer Centre Cyfronet AGH

- The largest Polish Academic Computer Centre
  - 50 years of experience in IT provision
  - Centre of excellence in HPC, Grid and Cloud Computing
  - Home for Athena, Ares and Prometheus supercomputers
  - LUMI consortium partner (EuroHPC pre-exascale supercomputer #3 on TOP500)
- Legal status: an autonomous within AGH University of Science and Technology
- Staff: 180+ , ca. 80 in R&D
- Leader of PLGrid: Polish Grid and Cloud Infrastructure for Science
- NGI Coordination in EGI e-Infrastructure



Academic Computer Centre Cyfronet AGH

Prometheus



- 2.40 PFLOPS
- 53 568 cores
- From 2015 to 2021 1st HPC system in Poland (475th on Top 500, 38th in 2015)

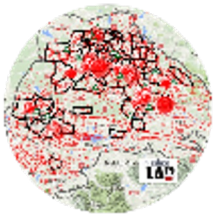
Athena

- 7.71 PFLOPS
- 384 A100 GPGPUs
- 1st HPC system in Poland (since 2022, 105th on Top500)
- 9th on Green500

Ares

- 4.00 PFLOPS
- 38 112 cores
- 290th on Top 500

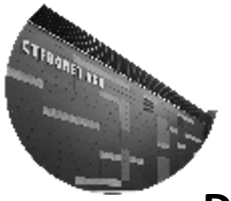
Computing portals and frameworks



- OneData
- PLG-Data
- Rimrock
- InSilicoLab

Storage

- 60+ PB
- hierarchical data management



Research & Development

- distributed computing environments
- computing acceleration
- machine learning
- software development & optimization

Data Centres

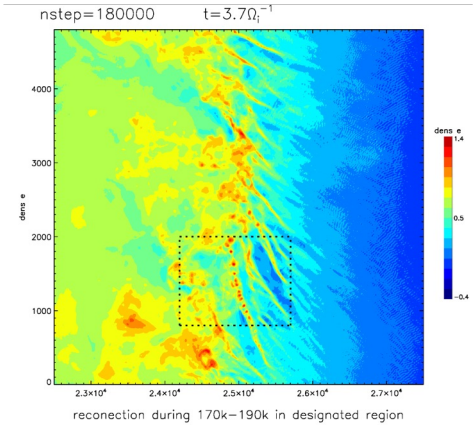
- 3 independent data centres
- dedicated backbone links

Computational Cloud

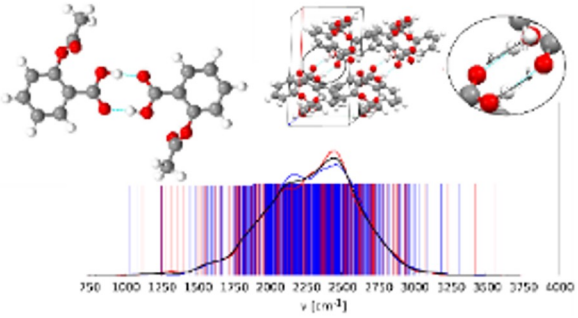
- based on OpenStack

Usage examples

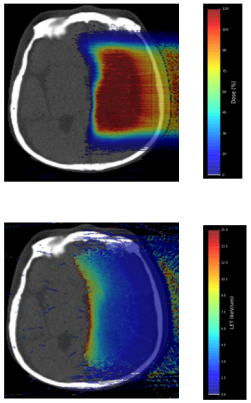
- Astrophysics: users' own code
  - Particle in Cell written in Fortran
  - Production runs - 9600 cores
    - One run up to 460 800 h CPU time – 53 years



- Chemistry: CPMD, CP2k, Jaguar, Gaussian
  - Importance of hydrogen bonds in biomolecules
  - Jobs: 24-240 cores
    - Hundreds of thousands of jobs with walltime < 1 h
    - Efficient usage resources through backfill

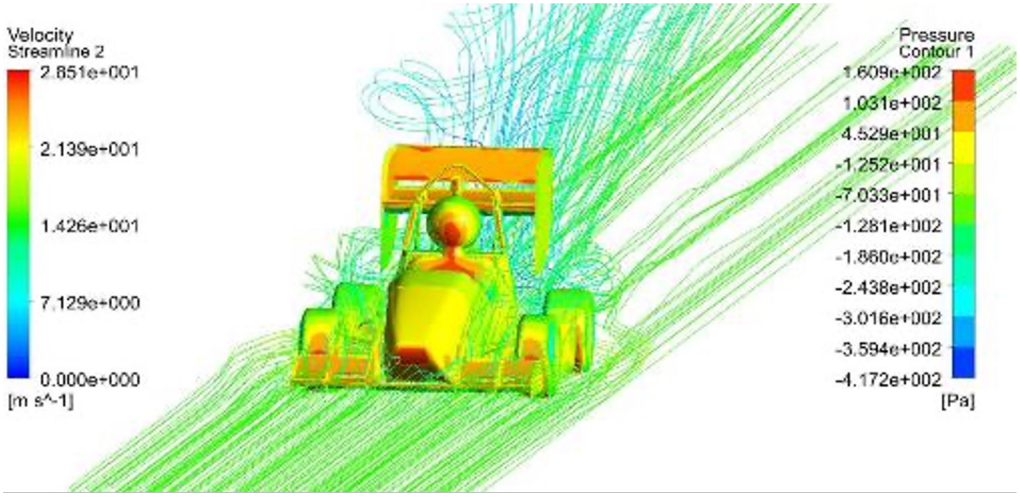


- Biophysics: Proton therapy
  - Monte Carlo simulations of a proton beam
  - Monte-Carlo based treatment planning
  - Jobs:
    - thousands of jobs with MC simulations (hours on hundreds of nodes)
    - Interactive large data processing with Jupyter noteboooks



Usage examples – AGH students scientific associations

- AGH Racing – Formula Student



- AGH Solarboat – Solar Sport One and Monaco Solar & Energy Boat Challenge

Open Publications

SKILLS
&
EXPERIENCE

Welcome to the PLGrid portal!

At this point you can prepare to use the PLGrid infrastructure - create affiliaiton, apply for a service or a grant. We encourage you to use PLGrid user manual.]
If it is necessary to contact Infrastructure employees, we recommend creating message in the system Helpdesk PLGrid.]

The most important functionalities

Affiliation

The PLGrid infrastructure provides resources to people associated with polish science. Enter the scientific unit under which you conduct the research. Then select whether you are an employee or if you work with some of employees (e.g. as a student or a PhD student).

Services

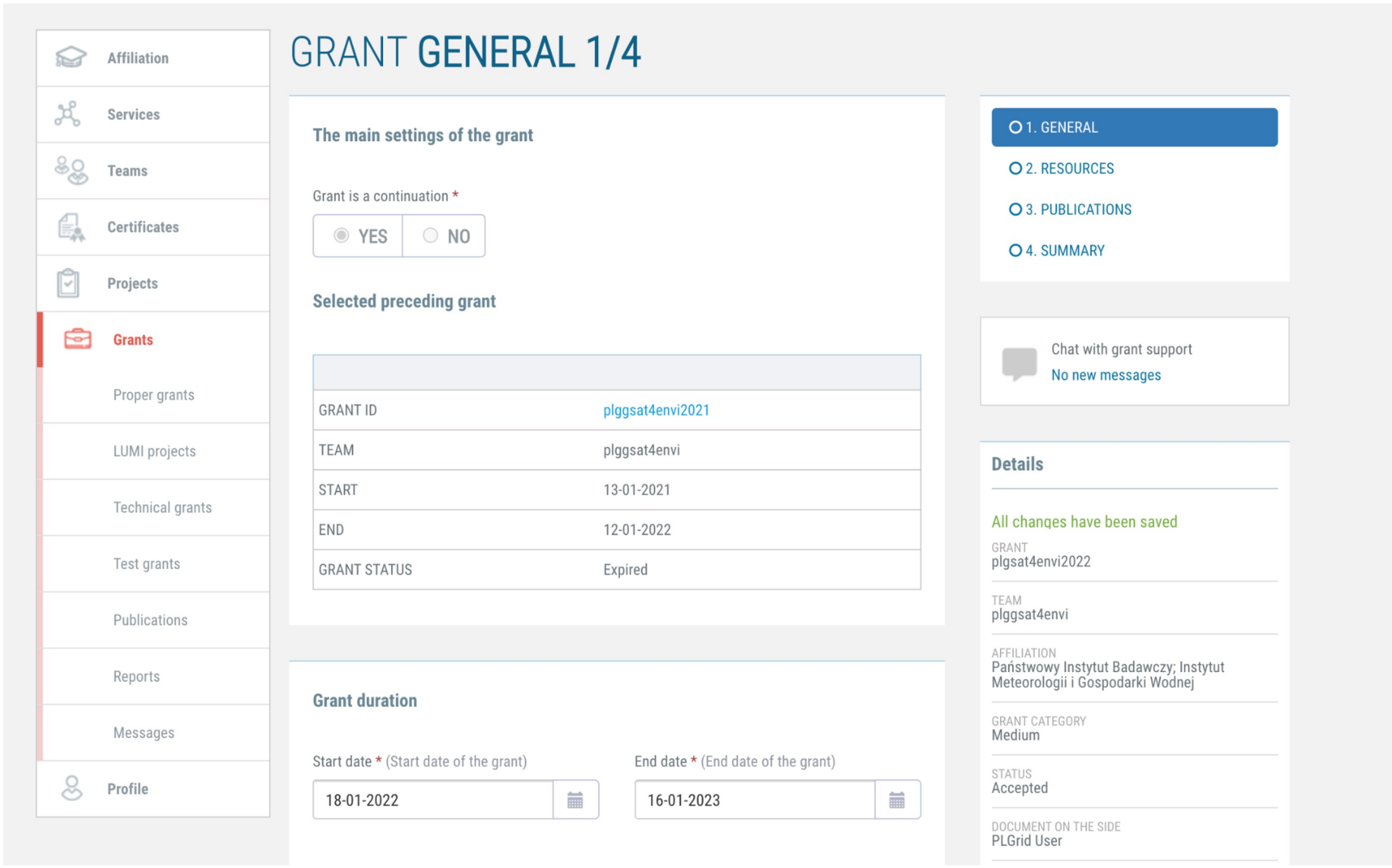
To be able to use the resources, select the services that interest you (big data, storage, cloud computing, social tools), ways to access them and additional services, which will improve the use of the Infrastructure.

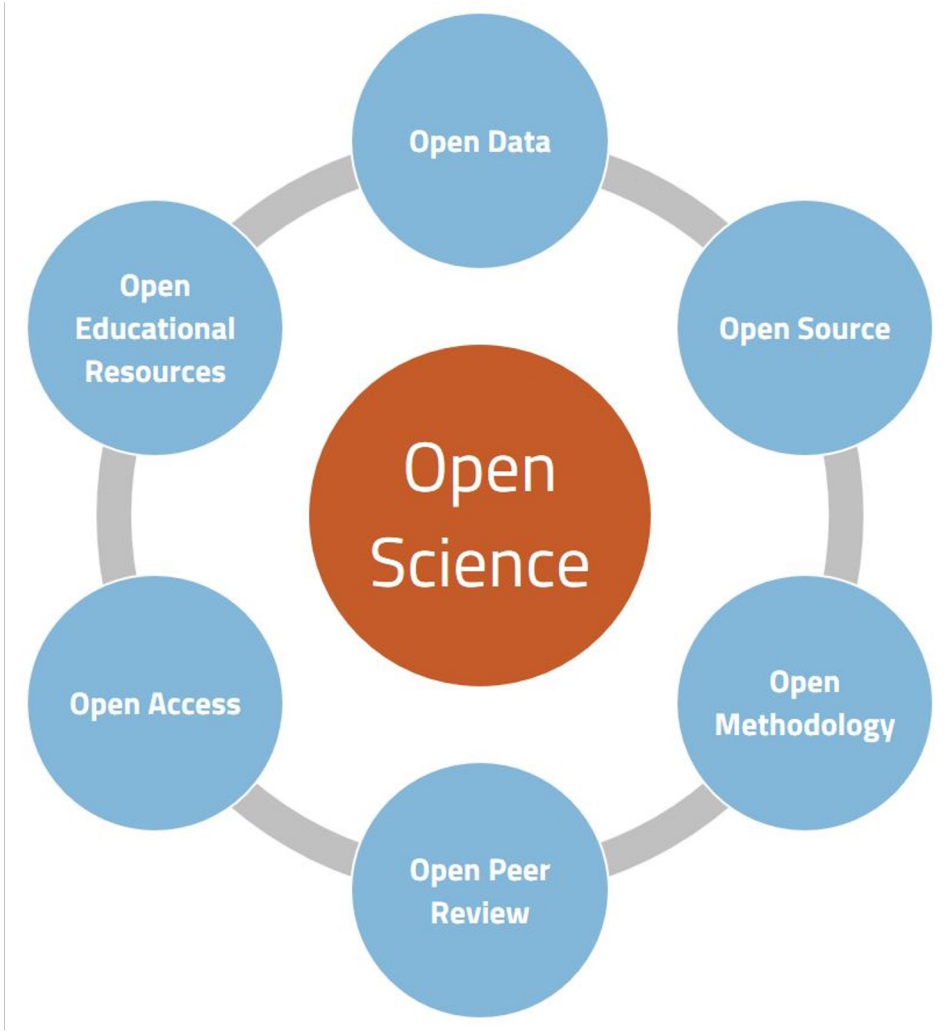
Teams

In order to request a proper grant, create a team that reflects the research group with which you want to share the grant. The team can be one person team.

Grants

The use of resources takes place within grants -contracts for the use of resources, defining the parameters of the resources made available. Apply for a grant, provide the research topics and technical parameters of the grant.

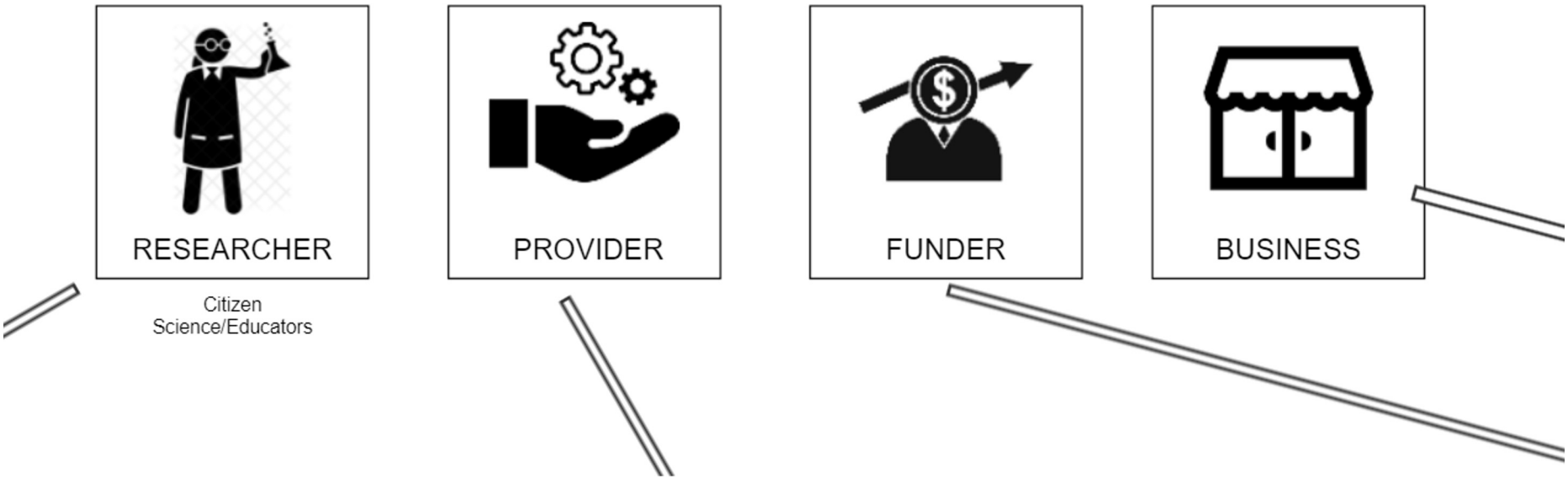

Affiliation
Services
Teams
Certificates
Projects
Grants
Proper grants
LUMI projects
Technical grants
Test grants
Publications
Reports
Messages
Profile
GRANT GENERAL 1/4
The main settings of the grant
Grant is a continuation *
YES
NO
Selected preceding grant
GRANT ID
plggsat4envi2021
TEAM
plggsat4envi
START
13-01-2021
END
12-01-2022
GRANT STATUS
Expired
Grant duration
Start date * (Start date of the grant)
18-01-2022
End date * (End date of the grant)
16-01-2023
1. GENERAL
2. RESOURCES
3. PUBLICATIONS
4. SUMMARY
Chat with grant support
No new messages
Details
All changes have been saved
GRANT
plgsat4envi2022
TEAM
plggsat4envi
AFFILIATION
Państwowy Instytut Badawczy; Instytut Meteorologii i Gospodarki Wodnej
GRANT CATEGORY
Medium
STATUS
Accepted
DOCUMENT ON THE SIDE
PLGrid User



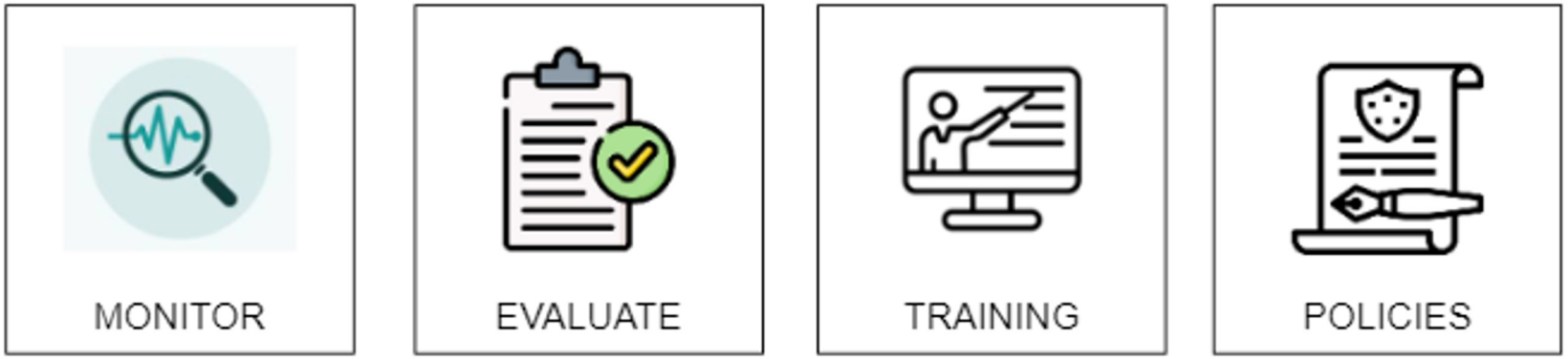
Source: https://www.openscience-maastricht.nl/the-open-science-mission/

Parties involved

WHO ARE YOU?
I AM A...

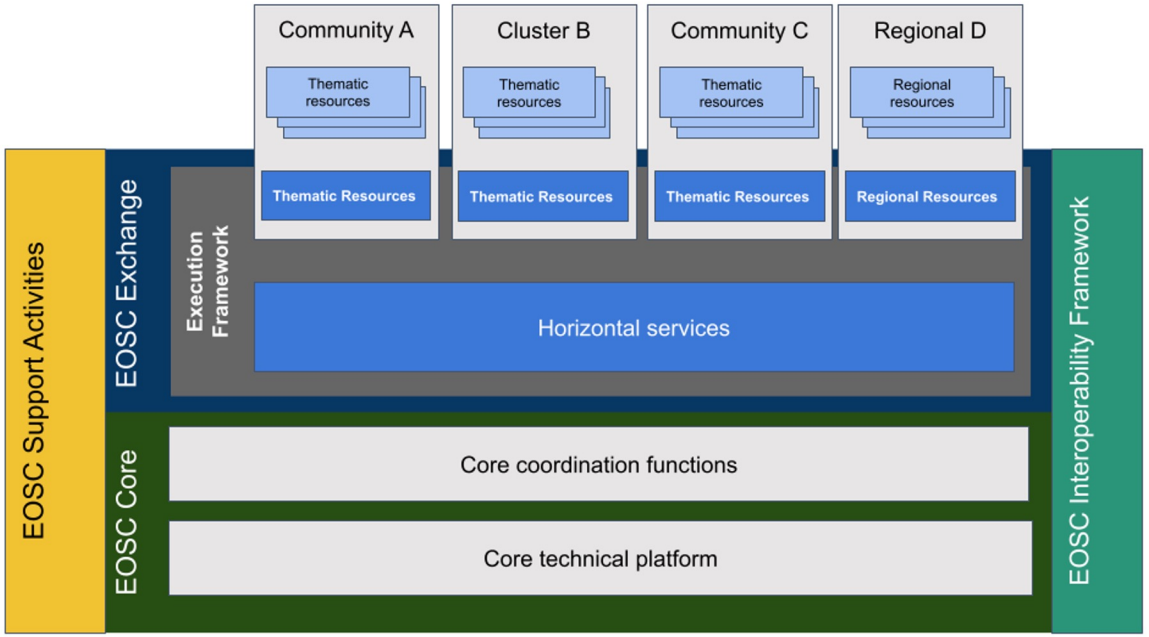


Funders

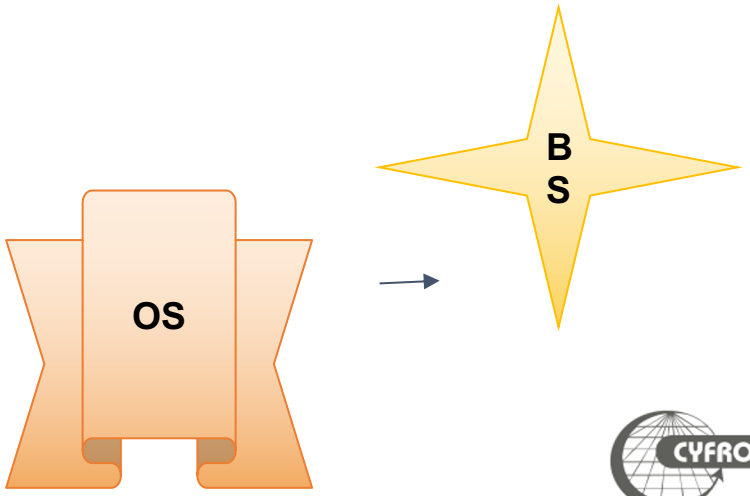


- Supporting tools →
- Features and operation ->
- **Knowledge and skills**

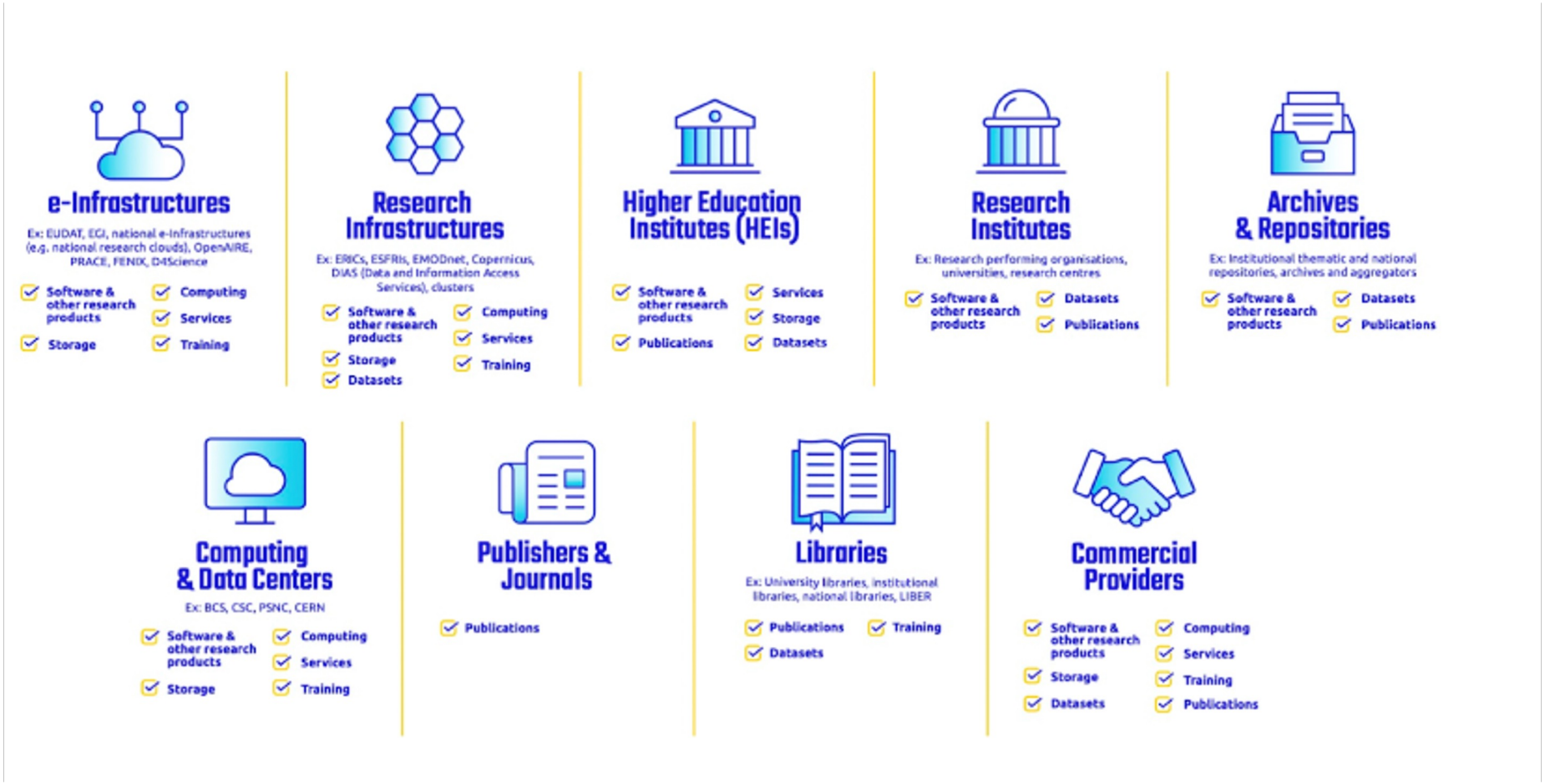
Providers



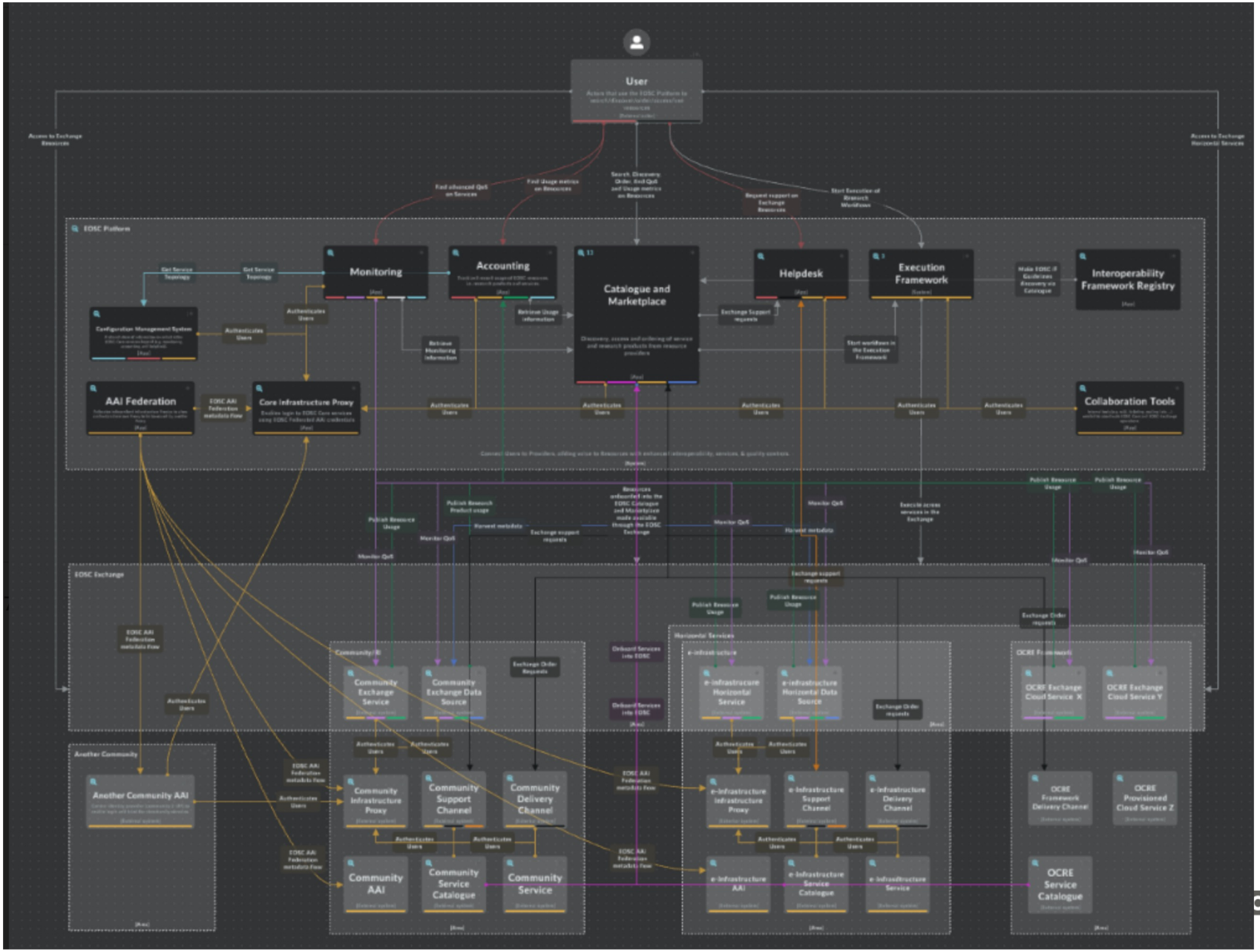
- Supporting tools →
- Features and operation ->
- **Knowledge and skills**



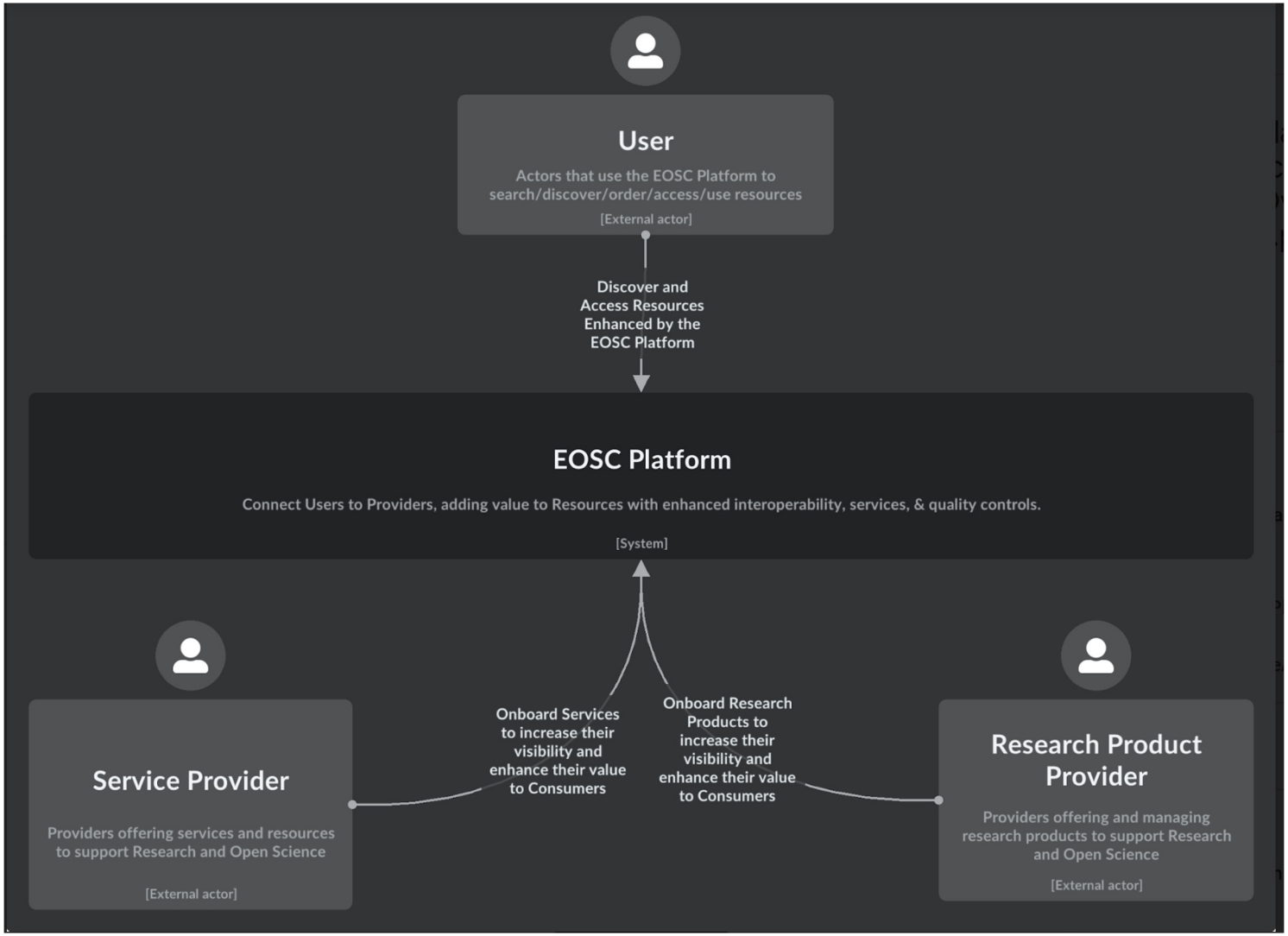
Providers one more time



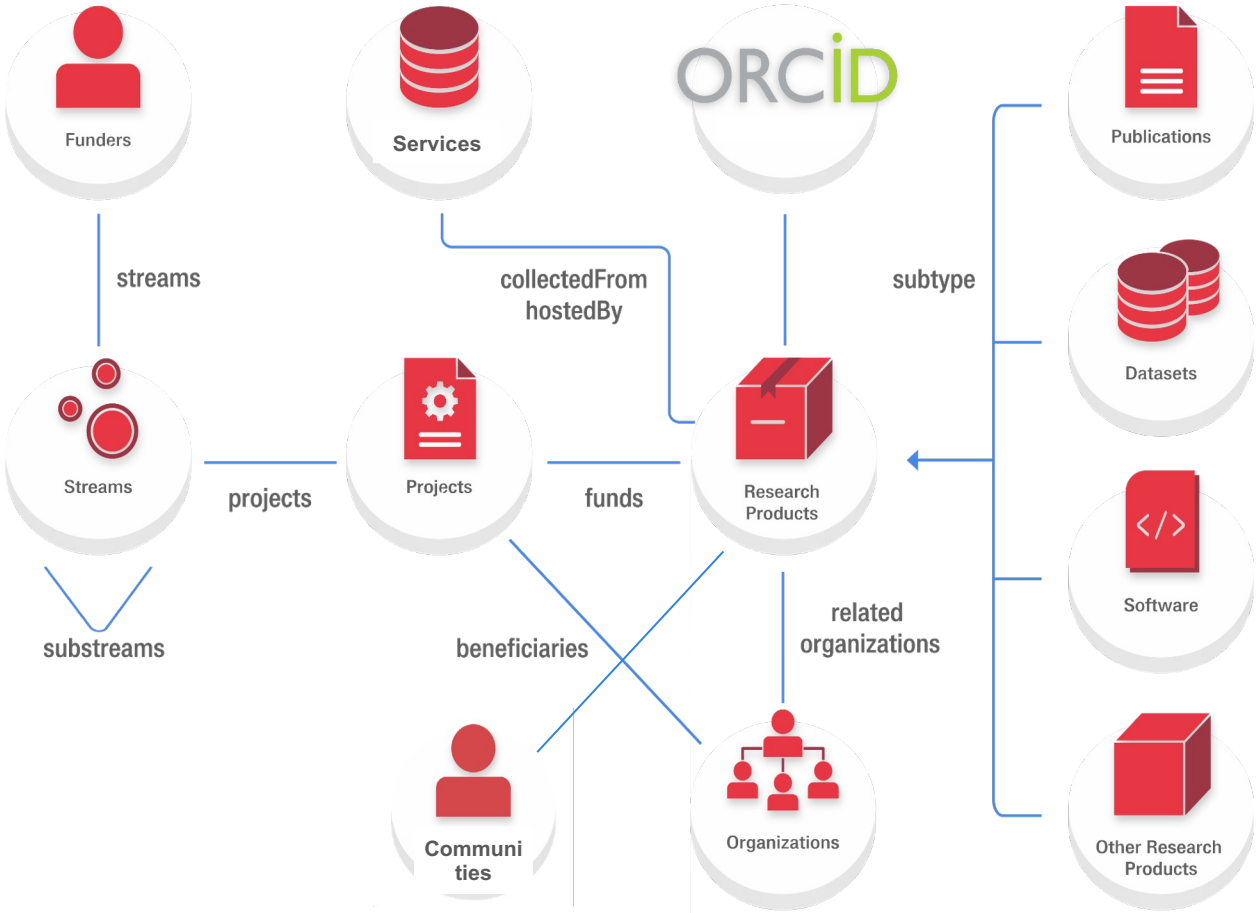
And again


User
EOSC Platform
Monitoring
Accounting
Catalogue and Marketplace
Helpdesk
Execution Framework
Interoperability Framework Registry
Configuration Management System
AAI Federation
Core Infrastructure Proxy
Collaboration Tools
EOSC Exchange
Community Exchange Service
Community Exchange Data Source
Another Community
Another Community AAI
Community Infrastructure Proxy
Community Support Channel
Community Delivery Channel
Community AAI
Community Service Catalogue
Community Service
e-Infrastructure Horizontal Service
e-Infrastructure Horizontal Data Source
e-Infrastructure Proxy
e-Infrastructure Support Channel
e-Infrastructure Delivery Channel
e-Infrastructure AAI
e-Infrastructure Service Catalogue
e-Infrastructure Service
OCRE Exchange Cloud Service X
OCRE Exchange Cloud Service Y
OCRE Framework Delivery Channel
OCRE Provisioned Cloud Service Z
OCRE Service Catalogue

Why, God, whyyyy??? Because…



Research products (EOSC RDGraph - Data model)



1. RAiD handle: a persistent, unique string of numbers, minted via the RAiD API.

2. The metadata envelope records time/date stamped PIDs for
   - Funders
   - Organisations
   - Collaborators
   - Tools and Services
   - Data

Source https://www.raid.org.au

Requirements from WP4

Research products

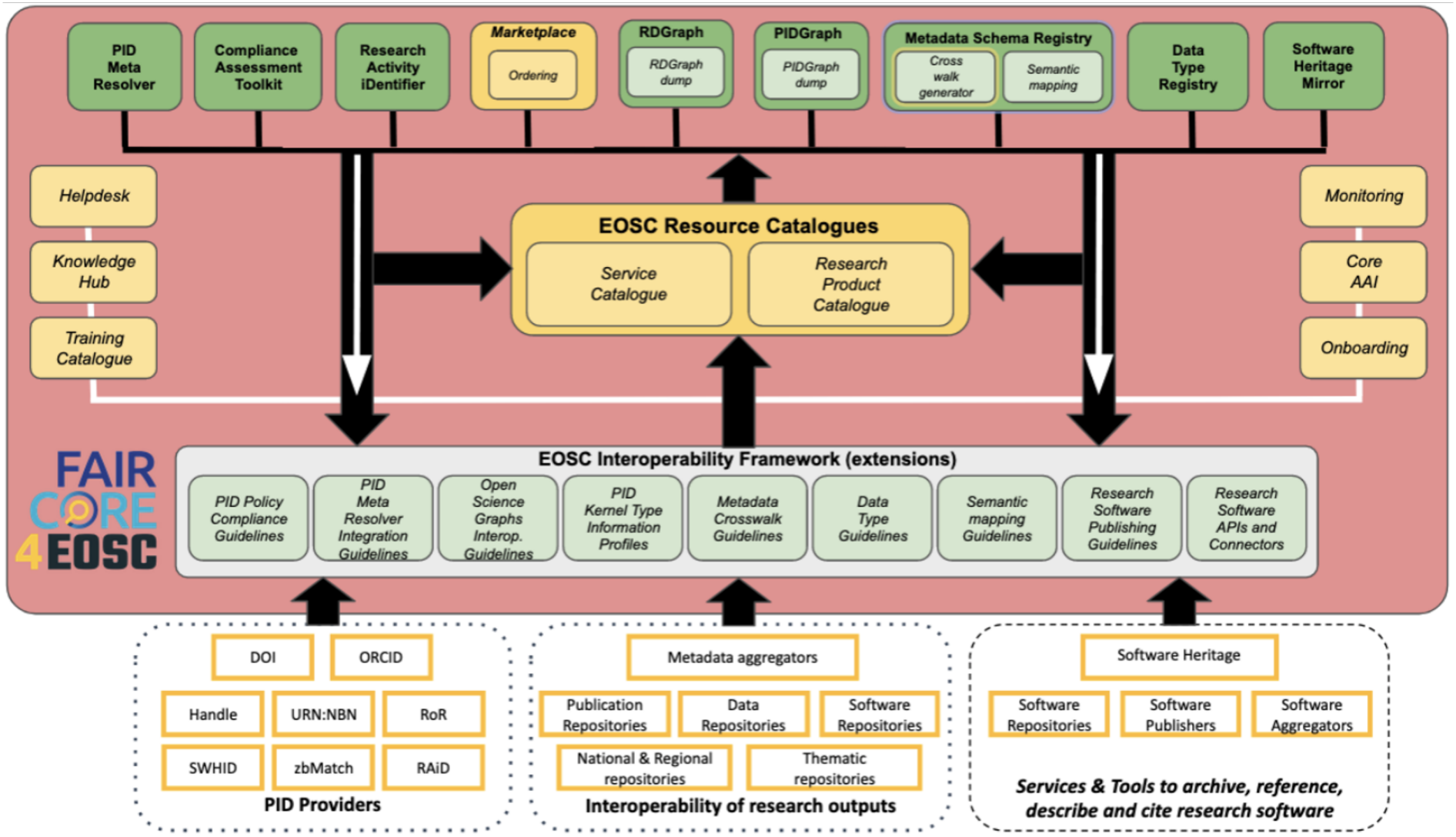
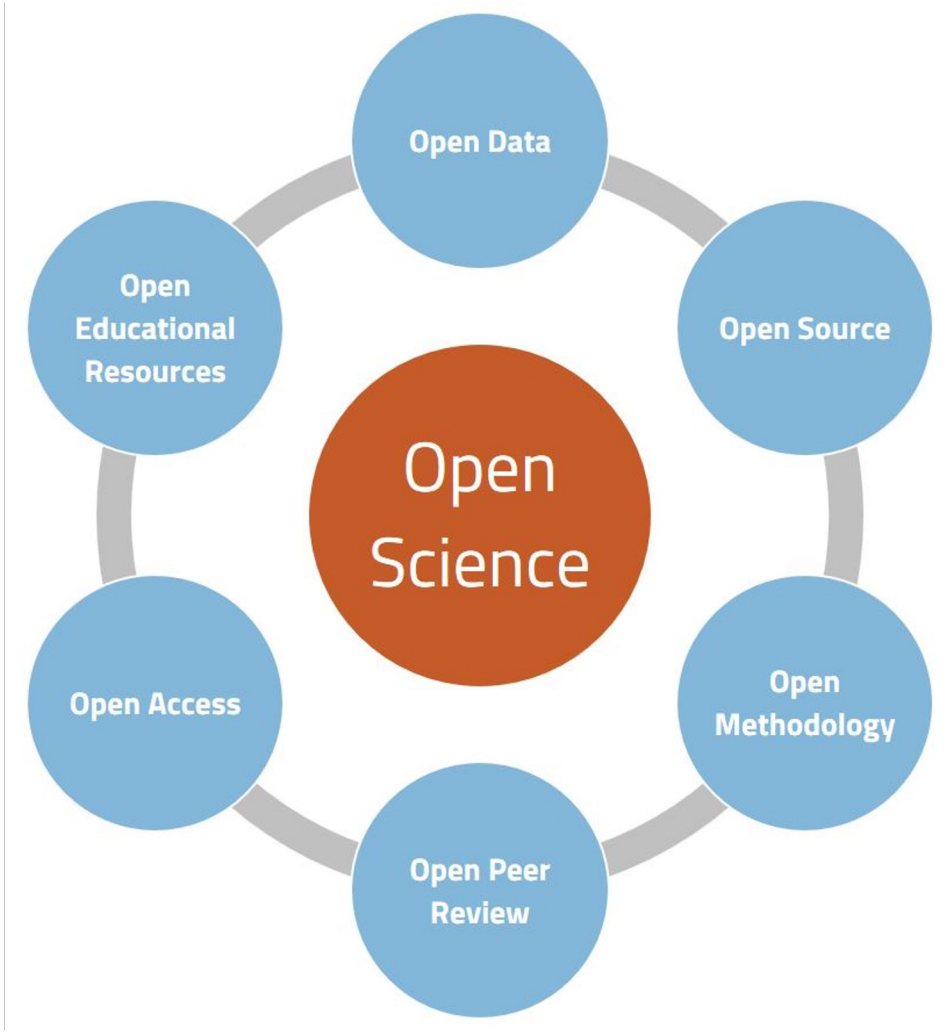
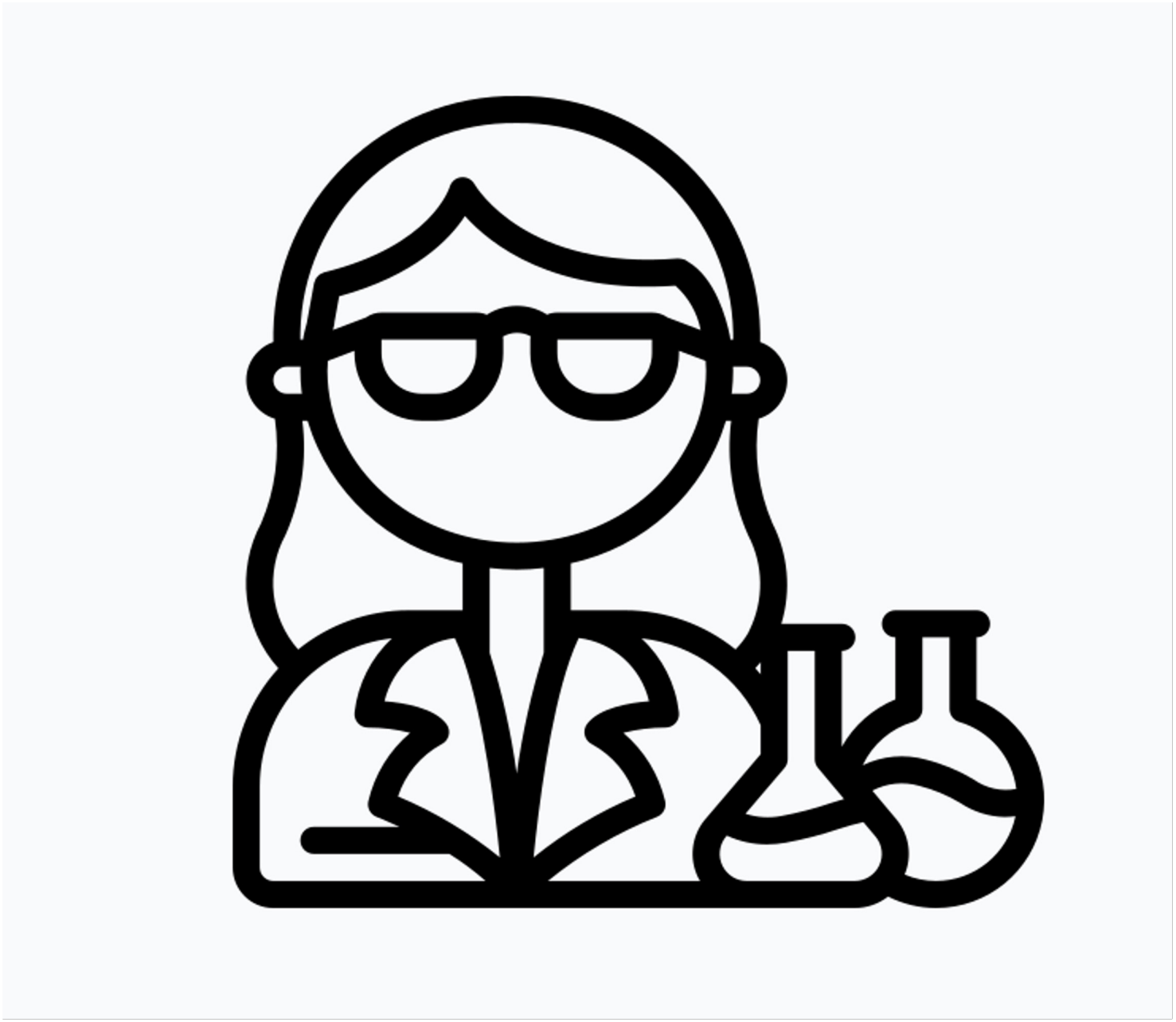


Figure 1 Architecture overview of the new FAIRCORE4EOSC components in relation to the EOSC-Core

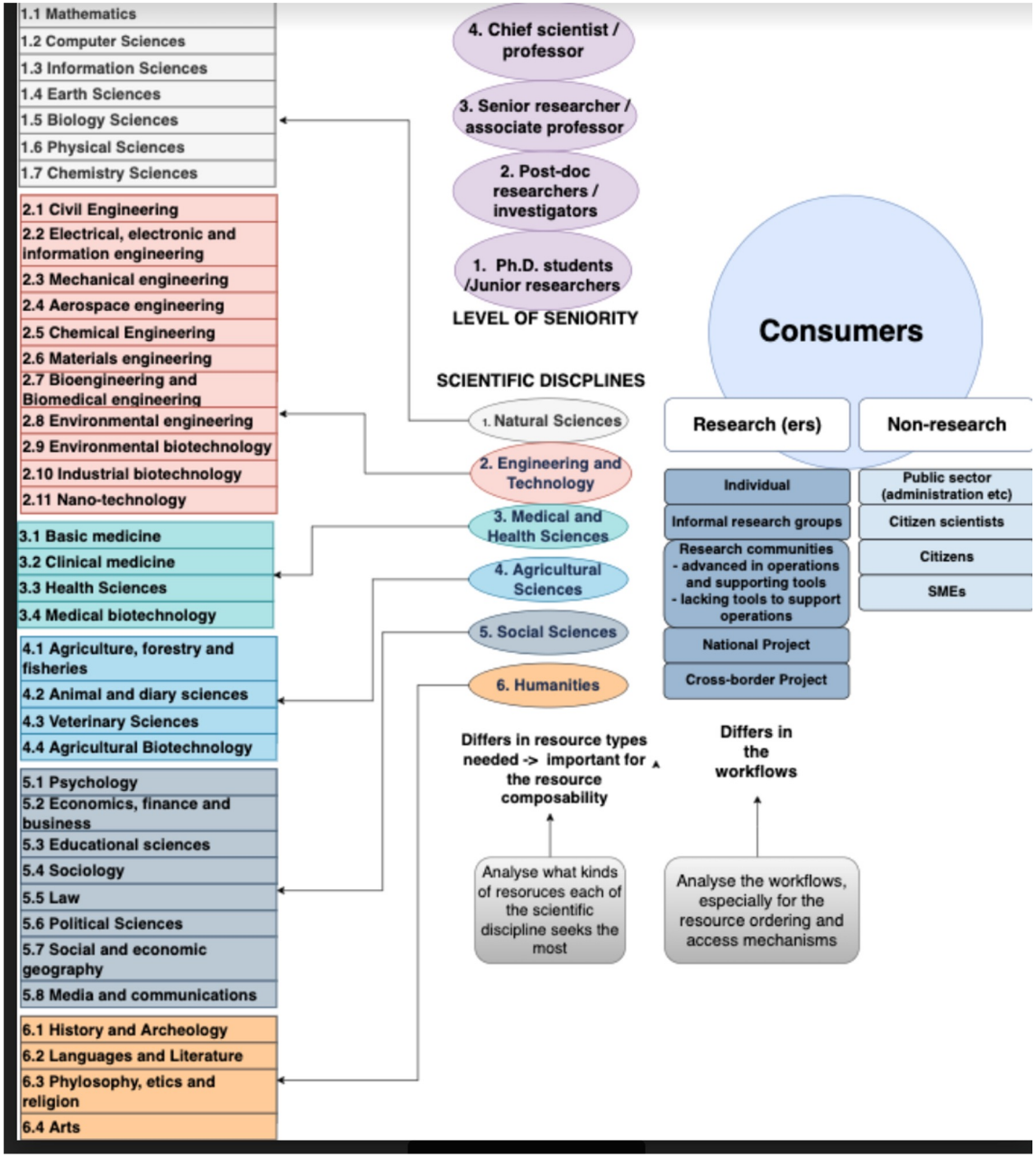


Source: https://www.openscience-maastricht.nl/the-open-science-mission/

And finally!



Consumers of Open Science



They come in as many flavours as you can imagine

Where should be the ‘know how’

SHARE
RESEARCH OUTPUTS

Store, backup, archive your data, publications, software

What
Data
Publications
Software

What to do
Store & Share
Backup
Archive

FIlters for data storage

- Type of storage
  - Generic
  - Field specific
- Size of dataset (MB)
- Assign DOI/PID (YES/NO)
- Security
- Price
  - Free
  - Fee
- Time span

It's not all…

- PROCESSING
  - Capabilities needed: transfer, compute (possibly external), **processing tool**, storage (possibly external)
  - What matters for this goal:
    - For storage:
    - Size of the dataset
    - Time of retention

- PUBLISHING YOUR DATASET
  - Capabilities needed: **publishing service**
  - What matters for this goal:
    - Size of the dataset
    - DOI
    - free/not free
    - Domain
    - Security (anonymise personal data/licence)

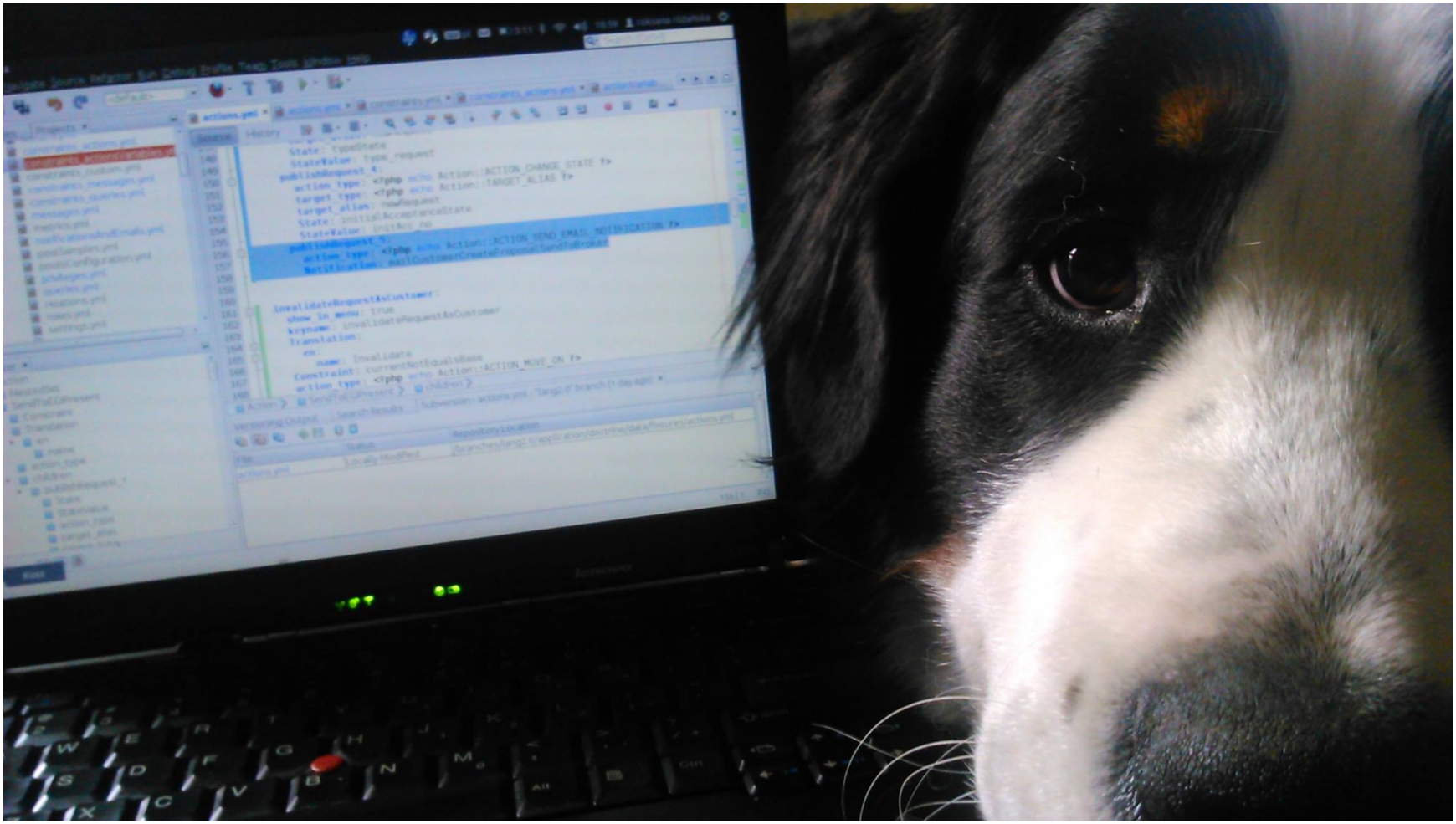
- ANONYMISATION

- MOVE/STORAGE YOUR (BIG) DATA (place where I can keep my data to use it during my processing)
  - Capabilities needed: transfer, **storage**

- USE EFFICIENT COMPUTING RESOURCES
  - Capabilities needed: **computing**, storage (possibly external)
- DISCOVER/RE-USE DATA
- MAKE YOUR REPOSITORY FINDABLE IN OPEN SCIENCE
  - Capabilities needed: **data sources**
  - What matters for this goal:
    - Connecting of metadata?
- **DATA MANAGEMENT**
- PRODUCE DATA (physical instruments)

CYFRONET

Thank you



r.wilk@cyfronet.pl