

# EOSC FESTIVAL 2023 National Tripartite Event Poland

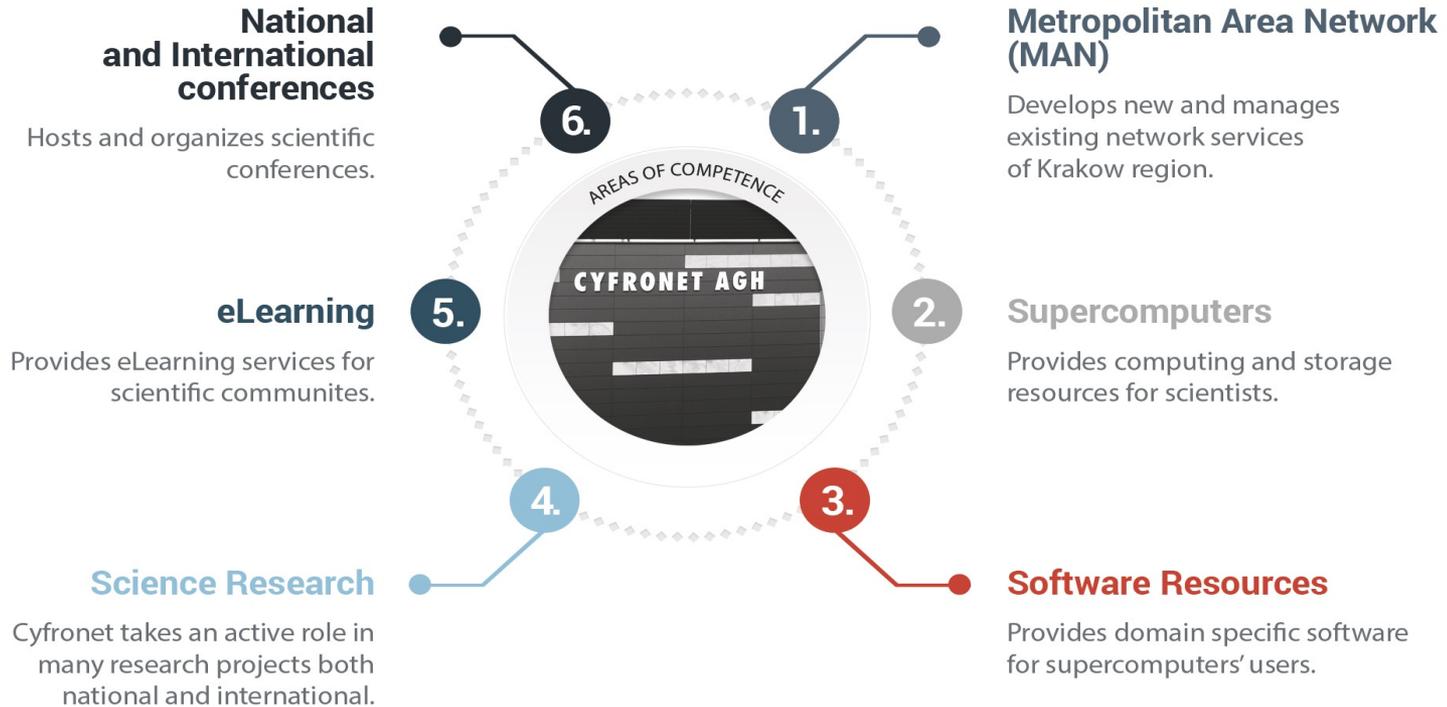
# Cyfronet's in EOSC and Open Science

Roksana Wilk

ACC Cyfronet AGH, Head of Data Processing Lab



# Cyfronet in a nutshell



# Cyfronet in a nutshell

- 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)
- Leader of PLGrid: Polish Grid and Cloud Infrastructure for Science
  - NGI Coordination in EGI e-Infrastructure
- Member of EOSC Association
- Active participant in EOSC Polish Network
- Certified against FitSM in service provision
- Cooperating with Research Data Alliance in data - related initiatives and projects

# Cyfronet and international initiatives



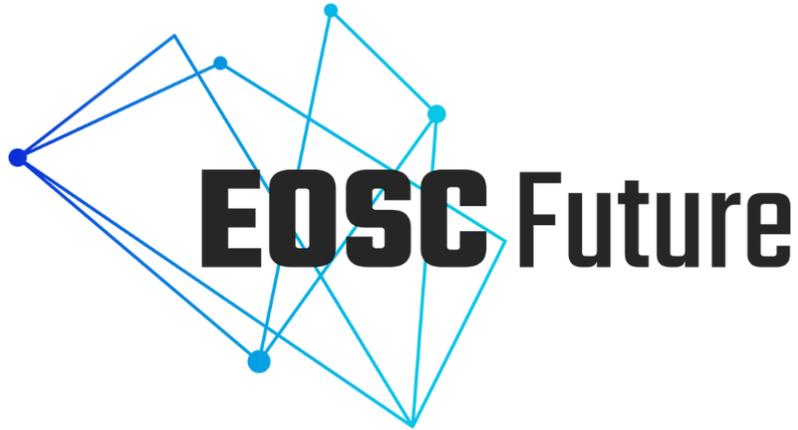
**EuroHPC**  
Joint Undertaking



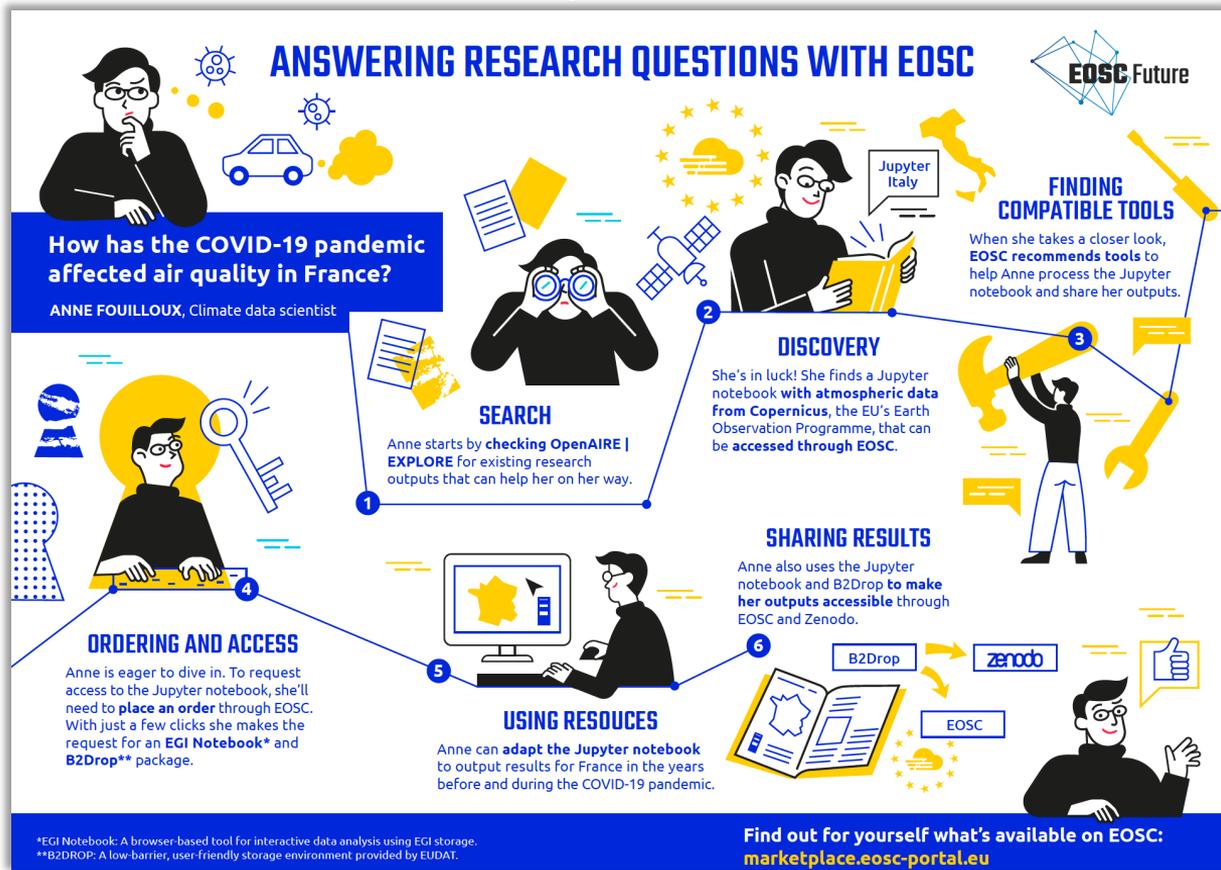
**sat4envi**



# Building EOSC



# Building the scientific ecosystem



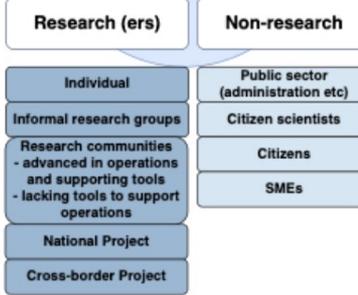
# Integrating Research-Performers



1.2 Computer Sciences
1.3 Information Sciences
1.4 Earth Sciences
1.5 Biology Sciences
1.6 Physical Sciences
1.7 Chemistry Sciences
2.1 Civil Engineering
2.2 Electrical, electronic and information engineering
2.3 Mechanical engineering
2.4 Aerospace engineering
2.5 Chemical Engineering
2.6 Materials engineering
2.7 Bioengineering and Biomedical engineering
2.8 Environmental engineering
2.9 Environmental biotechnology
2.10 Industrial biotechnology
2.11 Nano-technology
3.1 Basic medicine
3.2 Clinical medicine
3.3 Health Sciences
3.4 Medical biotechnology
4.1 Agriculture, forestry and fisheries
4.2 Animal and diary sciences
4.3 Veterinary Sciences
4.4 Agricultural Biotechnology
5.1 Psychology
5.2 Economics, finance and business
5.3 Educational sciences
5.4 Sociology
5.5 Law
5.6 Political Sciences
5.7 Social and economic geography
5.8 Media and communications
6.1 History and Archeology
6.2 Languages and Literature
6.3 Phylosophy, etics and religion

- LEVEL OF SENIORITY**
- 4. Chief scientist / professor
  - 3. Senior researcher / associate professor
  - 2. Post-doc researchers / investigators
  - 1. Ph.D. students / Junior researchers

- SCIENTIFIC DISCIPLINES**
- 1. Natural Sciences
  - 2. Engineering and Technology
  - 3. Medical and Health Sciences
  - 4. Agricultural Sciences
  - 5. Social Sciences
  - 6. Humanities



Differs in resource types needed -> important for the resource composability

Analyse what kinds of resources each of the scientific discipline seeks the most

Differs in the workflows

Analyse the workflows, especially for the resource ordering and access mechanisms



# And Research-Enablers



## e-Infrastructures

Ex: EUDAT, EGI, national e-Infrastructures (e.g. national research clouds), OpenAIRE, PRACE, FENIX, D4Science

- ✓ Software & other research products
- ✓ Storage
- ✓ Computing
- ✓ Services
- ✓ Training



## Research Infrastructures

Ex: ERICs, ESFRIs, EMODnet, Copernicus, DIAS (Data and Information Access Services), clusters

- ✓ Software & other research products
- ✓ Storage
- ✓ Datasets
- ✓ Computing
- ✓ Services
- ✓ Training



## Higher Education Institutes (HEIs)

- ✓ Software & other research products
- ✓ Publications
- ✓ Services
- ✓ Storage
- ✓ Datasets



## Research Institutes

Ex: Research performing organisations, universities, research centres

- ✓ Software & other research products
- ✓ Datasets
- ✓ Publications



## Archives & Repositories

Ex: Institutional thematic and national repositories, archives and aggregators

- ✓ Software & other research products
- ✓ Datasets
- ✓ Publications



## Computing & Data Centers

Ex: BCS, CSC, PSNC, CERN

- ✓ Software & other research products
- ✓ Storage
- ✓ Computing
- ✓ Services
- ✓ Training



## Publishers & Journals

- ✓ Publications



## Libraries

Ex: University libraries, institutional libraries, national libraries, LIBER

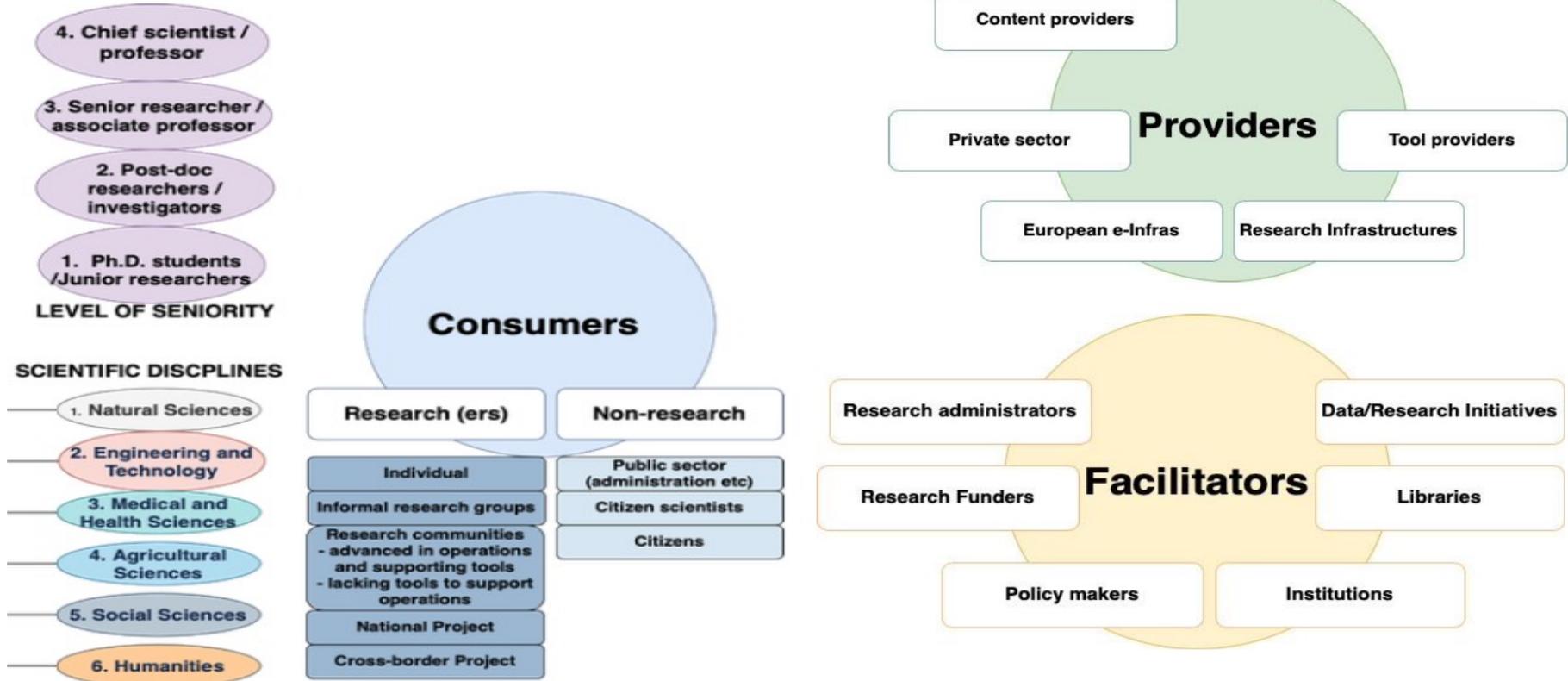
- ✓ Publications
- ✓ Datasets
- ✓ Training



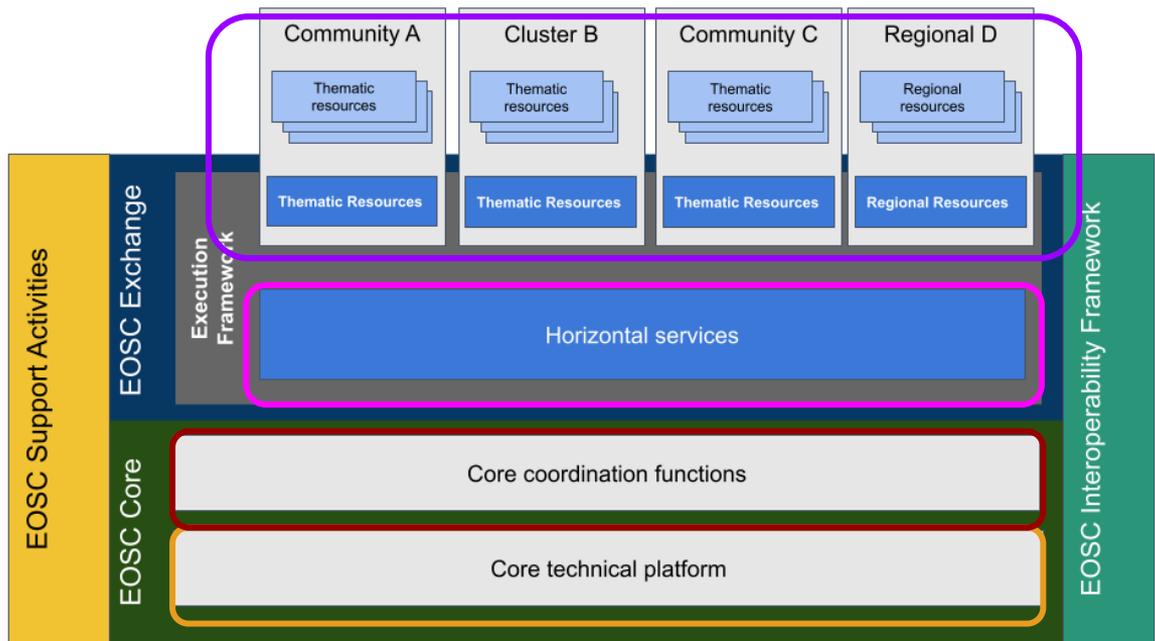
## Commercial Providers

- ✓ Software & other research products
- ✓ Storage
- ✓ Datasets
- ✓ Computing
- ✓ Services
- ✓ Training
- ✓ Publications

# EOSC user landscape



# The Architecture

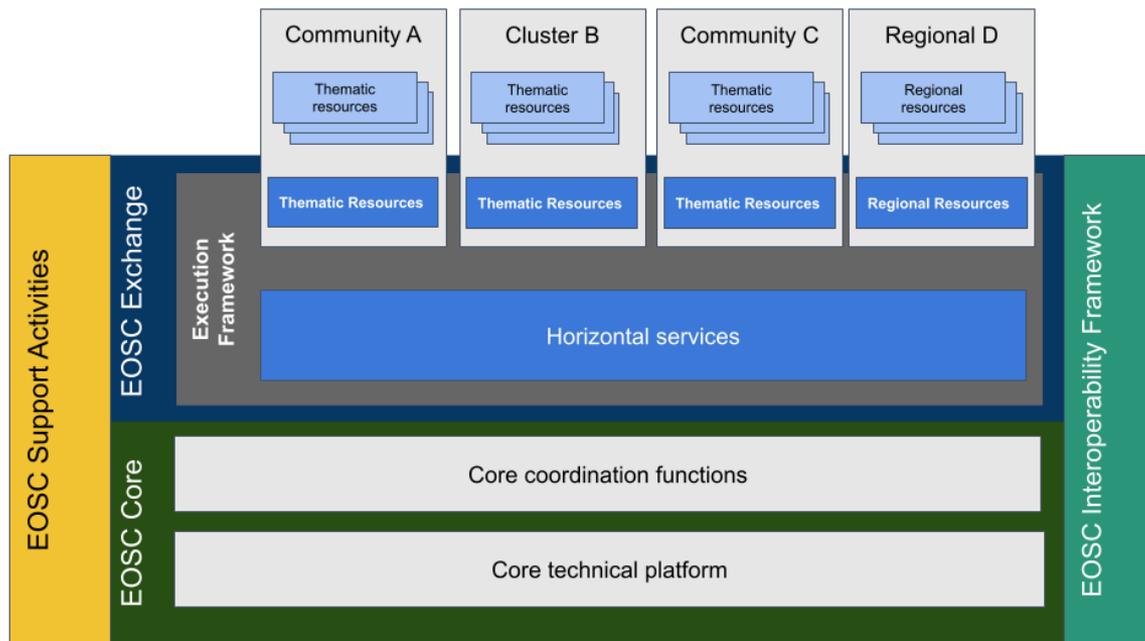


**Core:** Offering for research enablers  
Supporting services for professional and coordinated delivery of EOSC Exchange  
Operational layer of the technical CORE

**Exchange:** Offering for research  
General-use services able to add value to many scientific use case (e-Infra services)  
The thematic part of EOSC Exchange able to support a whole use cases coming from (one or more) scientific disciplines

**IF:** Composing environments for research  
Technical and operational layer enabling the composition of resources emerging from various scientific communities and providers

# The Architecture



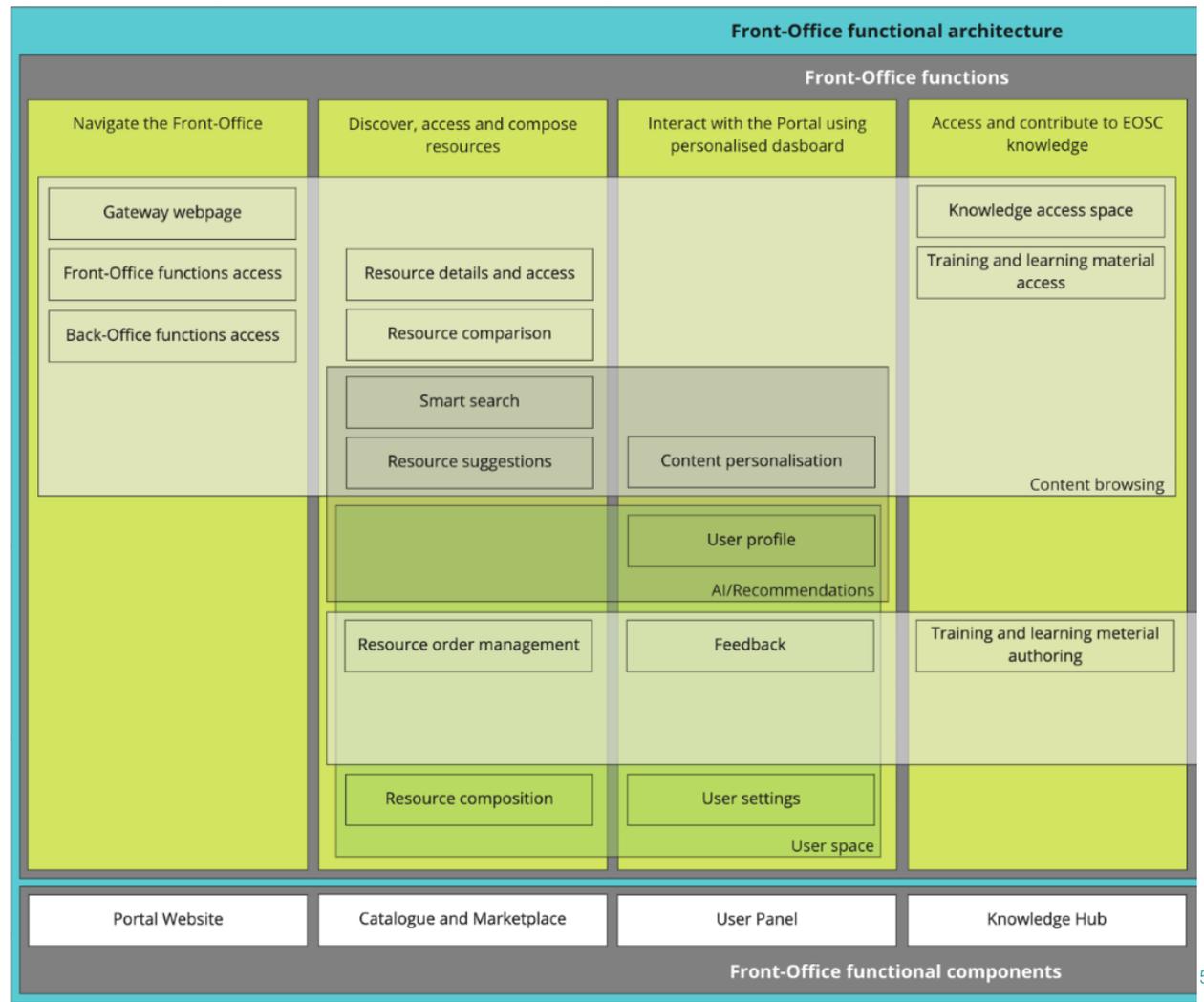
# EOSC Platform

EOSC Marketplace  
User Dashboard  
Knowledge Hub  
Order Management

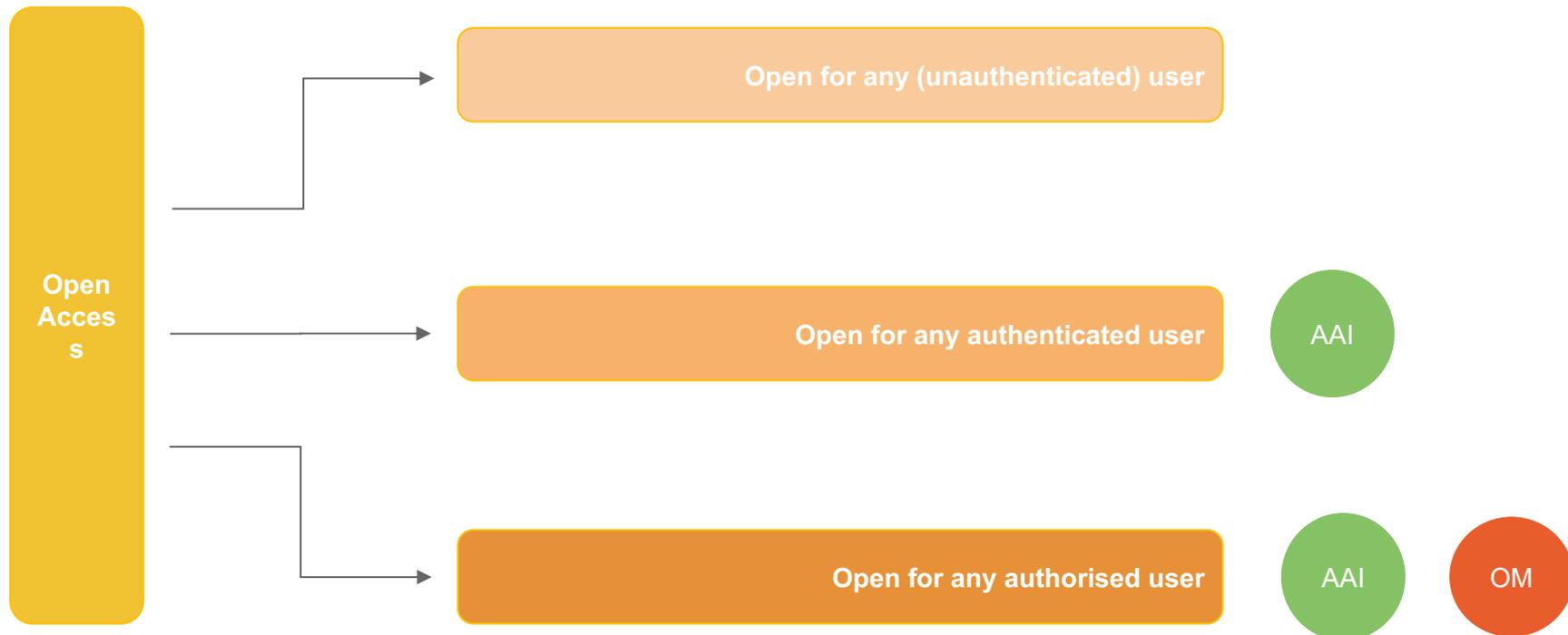
# EOSC Front Office



Set of components enabling access to user-facing services and functions of EOSC Node



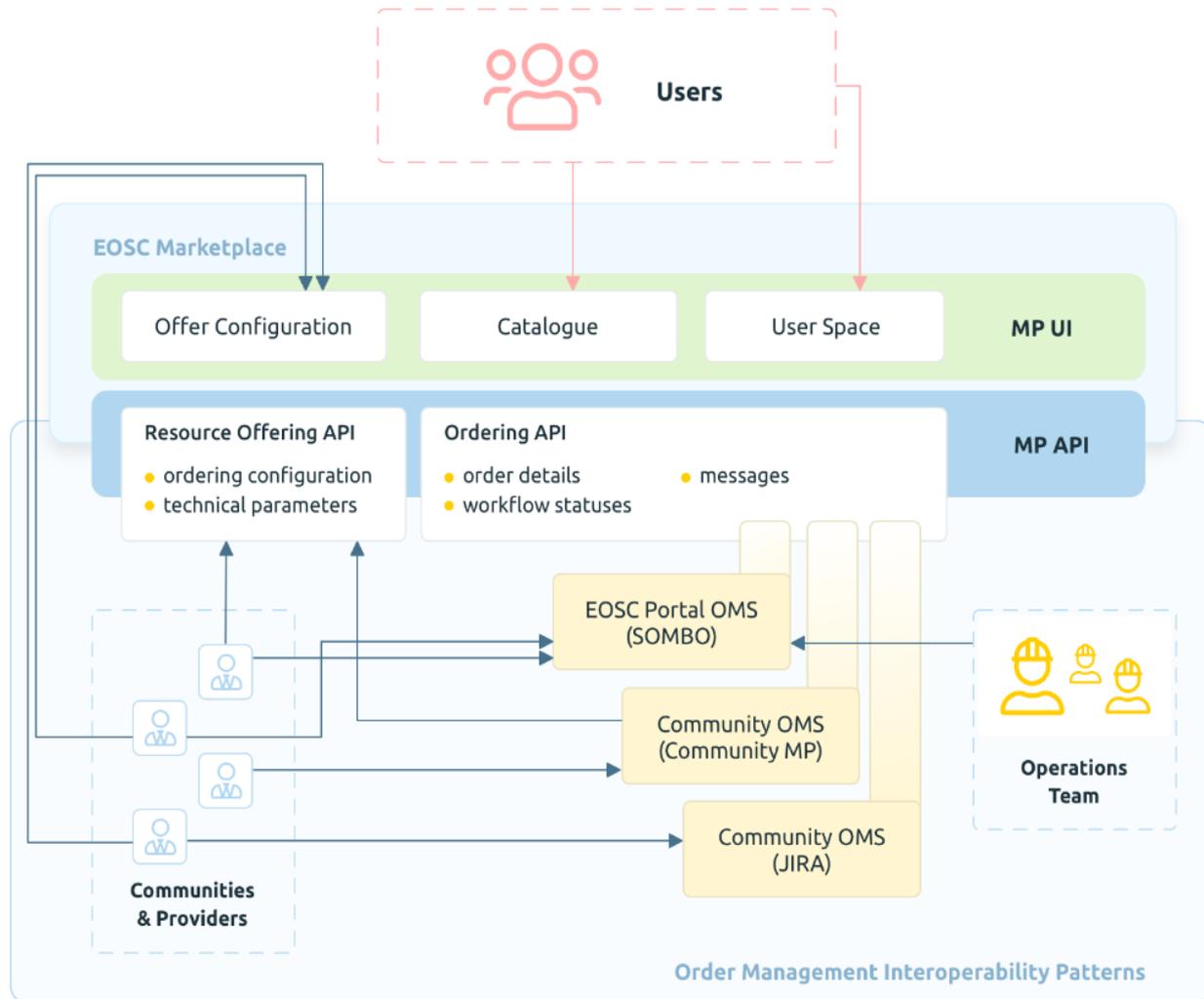
# Open access flavours in EOSC Marketplace



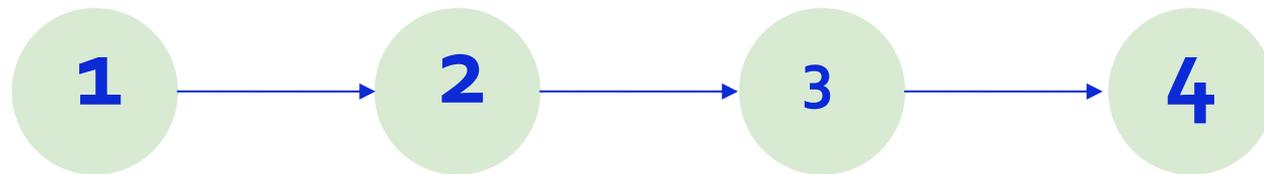
# Order management overview



Order Management is facilitating the **interaction between consumers and providers** that leads to **service utilization**.



# How it works



## Search Service

User can search the resources with various and specific filters to help them find the most suitable tools.

## Select an Offer

Choose among the various offerings of this service provided by the provider

## Choose parameters

Adjust the parameters of the requested offer to your needs

## Place and track the order

Track your order and contact your provider if you need

**After the Provider accepts the request, service is released to the user and user gets a notification**

# I found this interesting thing in EOSC

Research Data, Dataset, 2021  
**JUSThink Dialogue and Actions Corpus** ECI ANIMATAS

Norman, Utku; Dinkar, Tarvi; Nasir, Jauwairia; Bruno, Barbara; Clavel, Chloé; Dillenbourg, Pierre;

OPEN ACCESS ENGLISH

DOI: 10.5281/zenodo.4627104, 10.5281/zenodo.4627103  
Published: 22 Mar 2021

Download from [View less >](#)

ZENODO Dataset, 2021  
License: https://creativecommons.org/licenses/by/4.0/  
Providers: ZENODO

1

Data

Download

View all 3 versions  
Research Software, Software, 2022  
**JUSThink Alignment Analysis**

Norman, Utku; Dinkar, Tarvi; Bruno, Barbara; Clavel, Chloé;

OPEN ACCESS ENGLISH

Download from

ZENODO Software, 2022  
License: https://opensource.org/licenses/by/4.0/  
Providers: Datacite

2

Software

Download

**EGI Notebook**

EGI Notebook  
Create interactive documents with live code, visualisations and text  
Organisation: EGI Foundation  
Provided by: CESNET

☆☆☆☆☆ (0.0 / 5) 0 reviews  Add to comparison  Add to favourites

Access the resource

ORDER REQUIRED

3

Service

Access after  
Provider's OK

**Zenodo**

Zenodo  
A catch-all repository  
Organisation: OpenAIRE

OPEN ACCESS

Webpage Helpdesk Helpdesk e-mail User Manual  
Training information

4

Service

Access

Open  
Access



# EOSC Marketplace Projects

My Dashboard →

My Projects

- My scientific project1
- My scientific project2
- Sentinel Project

Create new project

## My scientific project1

Created at 12.08.2019 — Single user — ACC Cyfronet AGH

EDIT  
DUPLICATE

RESOURCES PROJECT DETAILS CONTACT WITH EOSC EXPERTS

### PROJECT DETAILS

<b>Project name</b> My scientific project1	<b>Email</b> r.wilk@cyfronet.pl
---	------------------------------------

### CUSTOMER DETAILS

<b>Customer Typology</b> Single user	<b>Origin country</b> Poland
<b>Organization</b> ACC Cyfronet AGH	<b>Webpage</b> http://www.cyfronet.pl

### USAGE

**Reason to request access to the EOSC services**

My very important research for the sake of exploring new things

My Dashboard →

My Projects

- My scientific project1
- My scientific project2
- Sentinel Project

Create new project

## My scientific project1

Created at 12.08.2019 — Single user — drf

EDIT  
DUPLICATE

RESOURCES PROJECT DETAILS CONTACT WITH EOSC EXPERTS

### Research products

**C++ & Python API for Scientific I/O with openPMD**

Software

Links ▾ | Unpin

**Call-by-Value and Call-by-Name Dual Calculi with Inductive and Coinductive Types**

Publication

Links ▾ | Unpin

### Ordered services

**EGI Cloud Compute** New

Service

### Open access services

**HADDOCK** Visit website

Service

**B2FIND** Visit website

Service

Add service to this project

# Next Goal

## → PROCESSING

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

## → PUBLISHING YOUR DATASET

- b) Capabilities needed: **publishing service**
- c) What matters for this goal:
  - ii) Size of the dataset
  - iii) DOI
  - iv) free/not free
  - v) Domain
  - vi) Security (anonymise personal data/licence)

## → DATA MANAGEMENT

## → 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

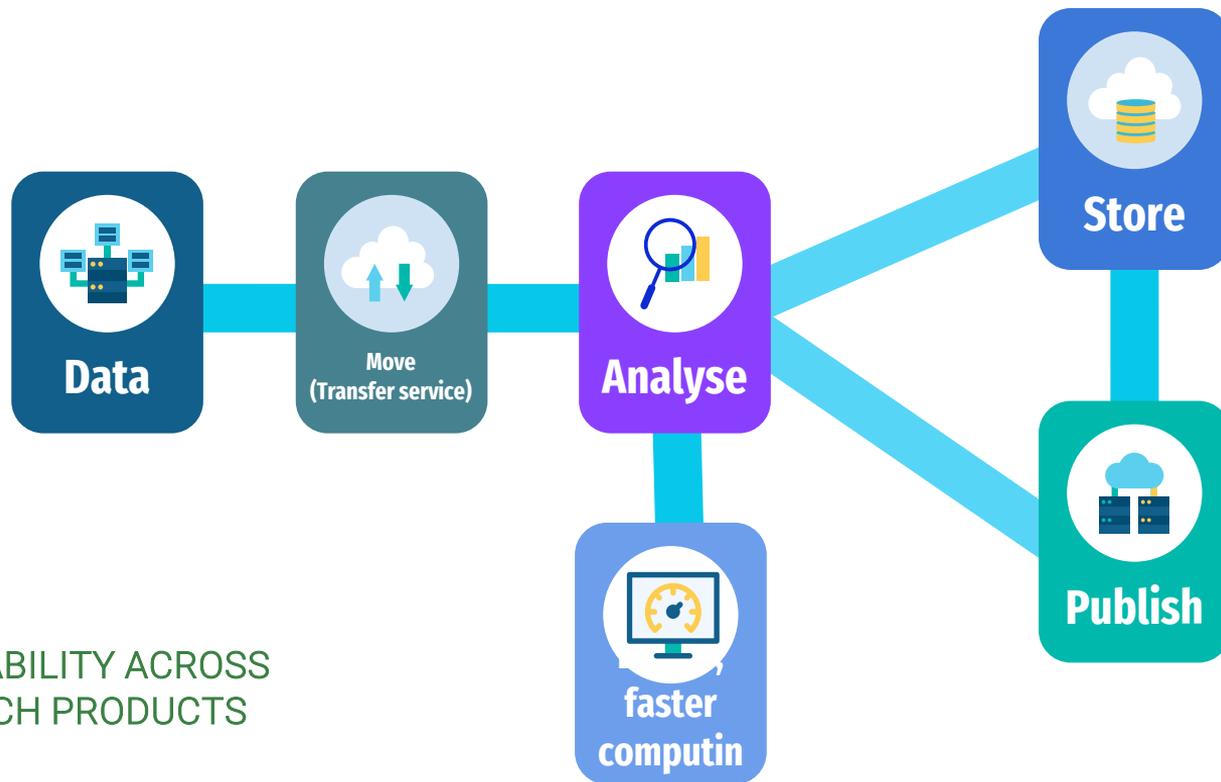
Capabilities needed: **data sources**

What matters for this goal:

- iii) Connecting of metadata?

## → PRODUCE DATA (physical instruments)

# Next Goal



INTEROPERABILITY ACROSS  
ALL RESEARCH PRODUCTS

# The Front-Office Service



<https://icecreamapps.com/v/d9x2gv2>

# EOSC Exchange

services for research

# PLGrid - (open) infrastructure for open science

- Open Access for recognised Polish researchers to services offered by major computing centres in Poland
- Centralised operations allowing easy access to computing, storage, cloud and thematic services
- Certified against FitSM standard
- Following the proposed architecture of EOSC Node and ready to integrate with its CORE services



<https://portal.plgrid.pl/>



# PLGrid - (open) infrastructure for open science

-  Affiliation
-  Services
-  Teams
-  Certificates
-  Projects
-  Grants
-  Subordinates
-  Profile

## Welcome to the PLGrid portal!

At this point you can prepare to use the PLGrid infrastructure - create affiliation, 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.

# PLGrid - (open) infrastructure for open science

The screenshot displays the PLGrid Portal interface. On the left is a dark sidebar with navigation options: Profile, Affiliations, Services, Research team, Certificates, Projects, Subordinates, and Grants (with a dropdown menu containing List, New Grant, and Settlements). The main content area features a search bar with fields for Title, ID, and Team, and a dropdown for All States. Below the search bar is a 'Sort by' dropdown. The main area lists four grants, each with a PLGrid logo, title, grant ID, team name, and status 'In the settlement'.

Title	ID	Team	Status	Last update
Udostępnianie cyfrowych zasobów nauki	plgsat4envi2022	plggsat4envi	In the settlement	2023-01-16 00:00
Gromadzenie danych satelitarnych	plggsat4envi2021	plggsat4envi	In the settlement	2022-11-16 10:23
Gromadzenie danych satelitarnych	sat4envi2020	plggsat4envi	In the settlement	2022-11-16 10:23
Gromadzenie danych satelitarnych	sat4envi2019	plggsat4envi	In the settlement	2022-11-16 10:23

# PLGrid - (open) infrastructure for open science

The screenshot displays the PLGrid Portal interface. On the left is a dark navigation sidebar with a hamburger menu icon at the top. The sidebar contains the following items: Profile, Affiliations, Services, Research team, Certificates, Projects, Subordinates, Grants (with a dropdown arrow), List, New Grant, and Settlements. The main content area is titled "PLGrid Portal" and "Grant". It features a "Basic information" section with fields for Title, Sort by, Grant title, Grant ID, Research team, and Affiliation. Below this are sections for Resources, Grant details, Related projects, Declared publications, Grant summary, and Settlement. A "MESSAGES" section shows a message from "System" with a status of "Accepted" and a timestamp of "2020-01-01 00:00". A "Previous grant where research was started" section links to "sat4envi2019". A table at the bottom lists grant details:

Title	Gromadzenie danych satelitarnych
Grant ID	sat4envi2020
Grant number	PLG/2019/013304
Research team	plggsat4envi
Affiliation	Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie / Akademiczne Centrum Komputerowe Cyfronet AGH
Grant type	PLGrid grant

# Rodbuk

- Access to all research products stored in Krakow's Universities' Libraries
- Allows to share and publish research products
- Implementing the FAIR principles
- Metadata schema aligned with EOSC requirements which supports research products' discovery and interoperability

## RODBUK



 **RODBUK** Cracow Open Research Data Repository

Data Source  OPEN ACCESS

↓ 0 Downloads    👁 62 Views    🌐 2 languages ▾

Organisation: Academic Computer Centre CYFRONET AGH

Scientific domain: Generic

The purpose of RODBUK is to collect develop archive and make available in open access all types of research data created by researchers PhD candidates and students in the course of scientific activity. RODBUK aims to implement the Open Science policy... [Show more](#)

[Access the Datasource](#)

<https://rodbuk.pl/>

# Rodbuk

## RODBUK

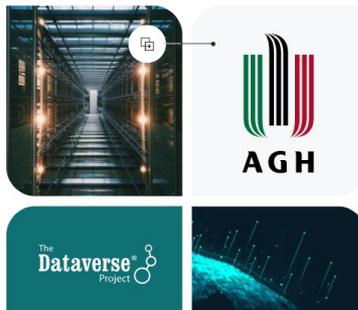
### RODBUK Cracow Open Research Data Repository

RODBUK was founded on the initiative of Board of Directors of the Cracow Library Group.

RODBUK is co-created by six Cracow universities:

- AGH University of Science and Technology,
- University of Physical Education in Krakow,
- Cracow University of Technology,
- Krakow University of Economics,
- Jagiellonian University in Kraków,
- University of the National Education Commission, Krakow.

User Guide Support English ▾



Search all dataverses...

Search

View all data >

### Deposit data in your Institutional Collection:



### RODBUK Cracow Open Research Data

RODBUK was founded on the initiative of Board of Directors of Cracow Open Research Data. RODBUK is co-created by six Cracow universities:

- AGH University of Science and Technology,
- University of Physical Education in Krakow,
- Cracow University of Technology,
- Krakow University of Economics,
- Jagiellonian University in Kraków,
- University of the National Education Commission, k

Search all dataverses...

### Deposit data



AGH University  
of Krakow



Jagiellonian  
University  
in Krakow

More...

Search this dataverse...



Advanced Search

**Dataverses (79)**

**Datasets (198)**

**Files (3,926)**

#### Dataverse Category

Department (58)

Organizational unit (21)

#### Publication Year

2023 (79)

#### Deposit Date

2023 (99)

2022 (1)

2021 (2)

2020 (3)

#### Branch of Science according to MEIN Domain/Subdomain

Natural sciences - biological sciences (46)

Natural sciences - chemical sciences (31)

Social sciences - psychology (24)

Natural sciences - physical sciences (20)

Medical and health sciences - medical sciences (11)

1 to 10 of 277 Results

Sort ▾

#### Replication Data for: Party-Related Primacy Effects in Proportional Representation Systems: Evidence from a Natural Experiment in Polish Local Elections



Nov 1, 2023 - [Jagiellonian University in Kraków](#)

Flis, Jaroslaw, Kaminski, Marek M., 2023, "Replication Data for: Party-Related Primacy Effects in Proportional Representation Systems: Evidence from a Natural Experiment in Polish Local Elections", <https://doi.org/10.57903/UJ/DS9GT4>

The dataset consists of parliamentary election results (party-level results and aggregate data about intra-party list distribution for major parties) in Polish general elections of 2011, 2015, and 2019. Territorially, data are aggregated at the level of Senate districts.

#### Jak Kochamy? Miłość romantyczna i kapitał kulturowy wśród miejskiej klasy średniej w indyjskich metropoliach. Baza materiałów z zakresu krytycznej analizy dyskursu - cz.1 (projekt nr 2020/37/B/HS3/04074)



Oct 31, 2023 - [Doctoral School in the Humanities](#)

Fołtyniak, Elzbieta; Anna Romanowicz, 2023, "Jak Kochamy? Miłość romantyczna i kapitał kulturowy wśród miejskiej klasy średniej w indyjskich metropoliach. Baza materiałów z zakresu krytycznej analizy dyskursu - cz. 1 (projekt nr 2020/37/B/HS3/04074)", <https://doi.org/10.57903/UJ/K5WJG6>

Dane badawcze dotyczące dyskursu miłosnego w Indiach zebrane w ramach projektu "Jak Kochamy? Miłość romantyczna i kapitał kulturowy wśród miejskiej klasy średniej w indyjskich metropoliach" (finansowanego przez Narodowe Centrum Nauki, nr 2020/37/B/HS3/04074). Research data on lov...

#### A set of measurement data from the ELA11 magnetometer from January 15, 2022, Hylaty Station, Poland



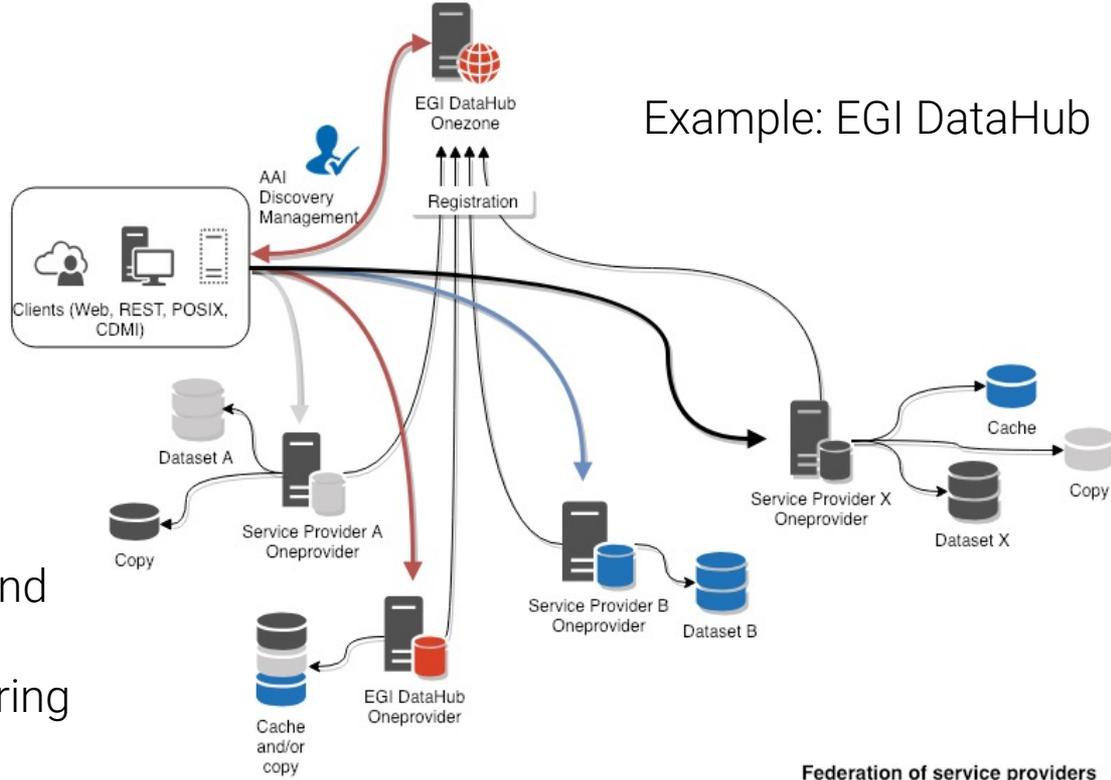
Oct 27, 2023 - [Faculty of Physics, Astronomy and Applied Computer Science](#)

Kubisz, Jerzy, 2023, "A set of measurement data from the ELA11 magnetometer from January 15, 2022, Hylaty Station, Poland", <https://doi.org/10.57903/UJ/NSRSLE>

A set of measurement ELF data from the ELA11 magnetometer from January 15, 2022, Hylaty Station, Poland

# ONE DATA distributed data management platform

- Unified access to geographically distributed data
- Ready for multi-cloud data access and processing
- Replica management with high-throughput data transfers
- Transparent data access under a single logical namespace
- Metadata management, indexing and data discovery
- Streamlined collaborative data sharing
- Comprehensive REST API for middleware integration



Example: EGI DataHub

# ONE DATA for OpenData

- Data collection (data set) aggregation service
- Easy sharing data for anonymous access
- Metadata indexing and data discovery
- Automation data processing workflows
- Shapshoting and curation of datasets
- Archivisation for long-term storage
- PID/DOI minting
- OAI-PMH for metadata harvesting

The screenshot shows the ONE DATA interface. On the left is a vertical navigation menu with icons for Data, Shared, Spaces, Groups, Tokens, and Providers. The main content area displays the details for a dataset named 'OxfordFlowerDatabase-FlowerSet1'. It includes the path, a public URL, and tabs for 'BASIC', 'JSON', and 'RDF'. Under the 'BASIC' tab, there is a 'license' field set to 'CC-0' and an 'Attribute' field. Below this are buttons for 'Save all changes', 'Discard changes', and 'Remove metadata'. A table below shows the files in the dataset:

FILES	SIZE
image_0001.jpg	50.85 KB
image_0002.jpg	41.24 KB
image_0003.jpg	46.2 KB
image_0004.jpg	30.16 KB

The screenshot shows a workflow visualization with three parallel processes. Each process is represented by a 'Parallel box' containing a task icon and name. The tasks are 'validate', 'unpack-data', and 'unpack-fetch' for the first process; 'unpack' and 'unpack-fetch' for the second; and 'fetch-files' for the third. Above the workflow, a status bar indicates 'Status: Finished'. Navigation arrows and a '1' in a green circle are visible between the process boxes.

# Sat4envi

- Enabling satellite data coming from the Sentinel satellites of the Copernicus network
- Supporting FAIR principles
- introducing a dedicated metadata schema empowering better findability
- dedicated for researchers and citizens
- dedicated part of the platform to support work of Polish administration and crisis management cells



<https://sat4envi.imgw.pl/>

# Sat4envi

ZDJĘCIA SATELITARNE SZUKAJ DANYCH

Wpisz szukaną miejscowość...

**POWIERZCHNIA**

- Charakterystyka powierzchni o wysokiej rozdzielczości

**POWIETRZE**

- Pył w atmosferze

**ATMOSFERA/METEOROLOGIA**

- Zachmurzenie (10.8µm)
- Chmury konwekcyjne

**DESZCZE I BURZE**

- Obraz satelitarny z rozdwojenymi

**WARSTWY WMS**

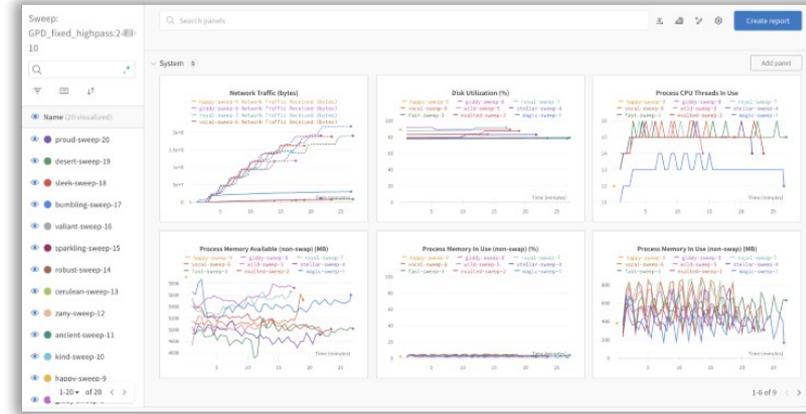
- województwa
- powiaty
- gminy
- jednostki\_ewidencyjne
- obreby\_ewidencyjne

przerwy serwisowej, dane satelitarne nie będą się aktualizowały w okresie od 31.07.2023 do 08.08.2023. Dane pobrane do serwisu po zakończeniu prac. Za utrudnienia przepaszamy.



# EPOS AI Platform

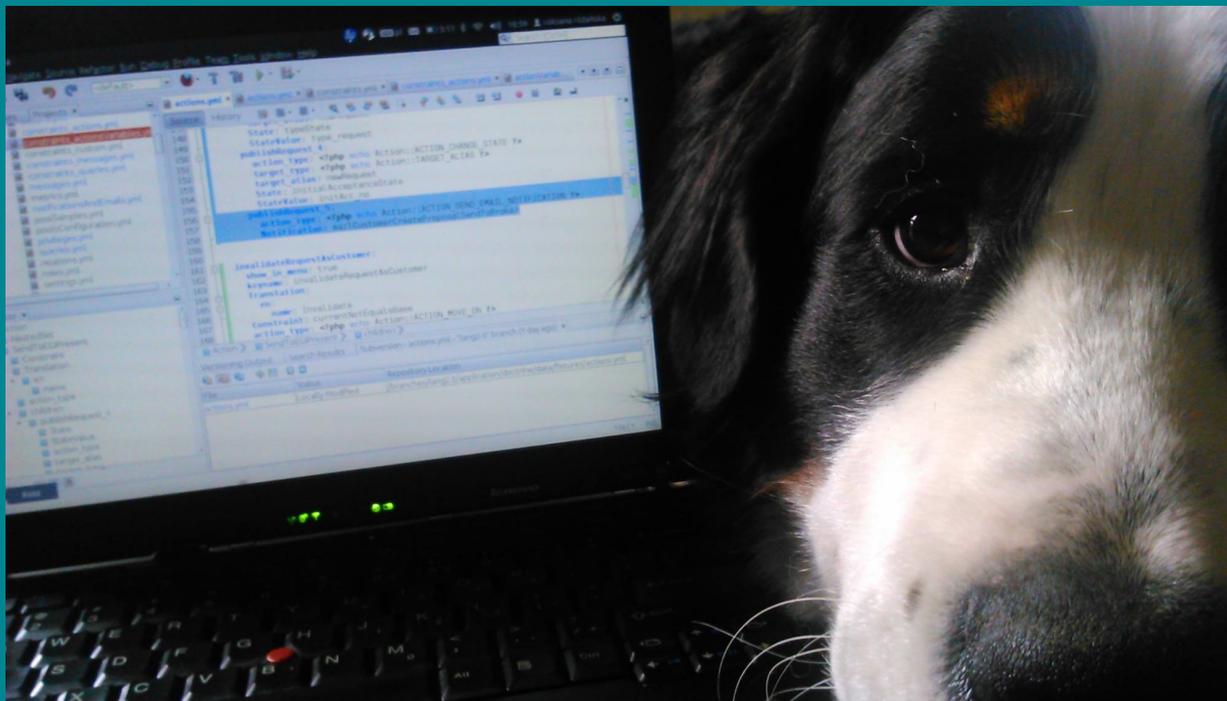
- Framework / tools facilitating creating, training and sharing ML models
- Connected to the HPC resources
- ACC Cyfronet AGH / PLGrid HPC resources
- Connected to the EPISODES Platform



<https://tcs.ah-epos.eu/>

# Other activities

- Co-chairing EOSC Rules-of-Participation Task Force
- Participation in FAIR Metrics and Data Quality Task Force
- Participation in starting CoARA's Ethics and Research Integrity Policy in Responsible Research Assessment for Data and Artificial Intelligence (ERIP-WG)
- Participation in policy-related activities led by CERN-NASA open-science collaboration
- Leading a working group of local Universities Libraries towards better metadata quality and defining metadata schemas of research products for the sake of research infrastructures and their capabilities
- Together with SANO investigating and implementing use cases of opening up (as open as possible as closed as necessary) sensitive medical data
- Activity in groups investigating the relation and integration between EOSC and EURO HPC



Thank you!  
r.wilk@cyfronet.pl