Emilie Hermans
EOSC-Pillar, Open Science
Ghent University,
Belgium

EOSC-Pillar

Coordination and Harmonisation of National &Thematic Initiatives to support EOSC

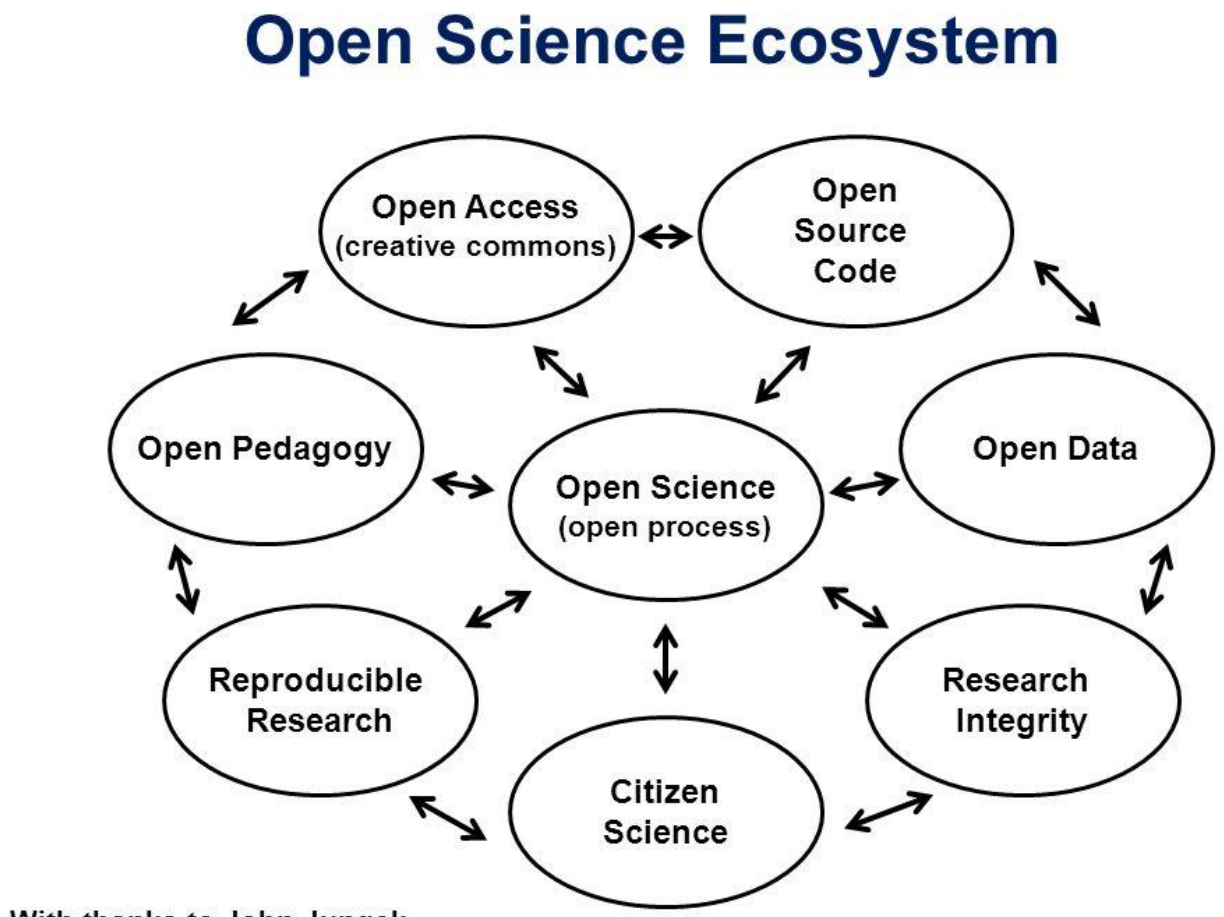
Open and FAIR data management

Open and FAIR data practices
for researchers and services, FAIR-enabling and the EOSC market place

EOSC-Pillar has received funding from the European Union's Horizon 2020 research and innovation Programme under Grant Agreement No. 857650.

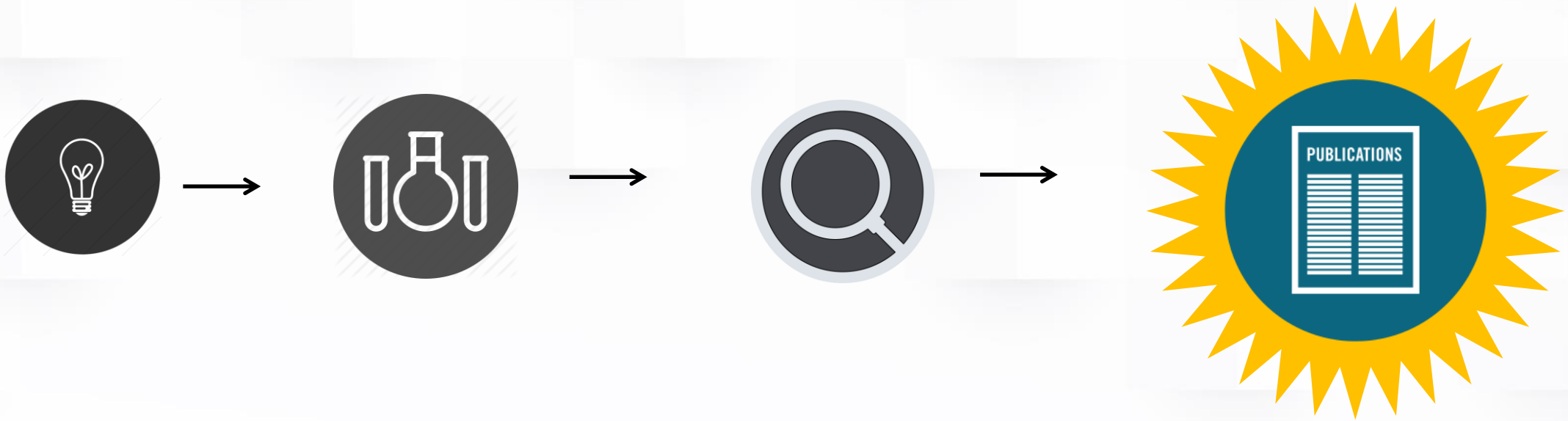
Some context

Open data and data management is part of a bigger movement towards more transparency in research



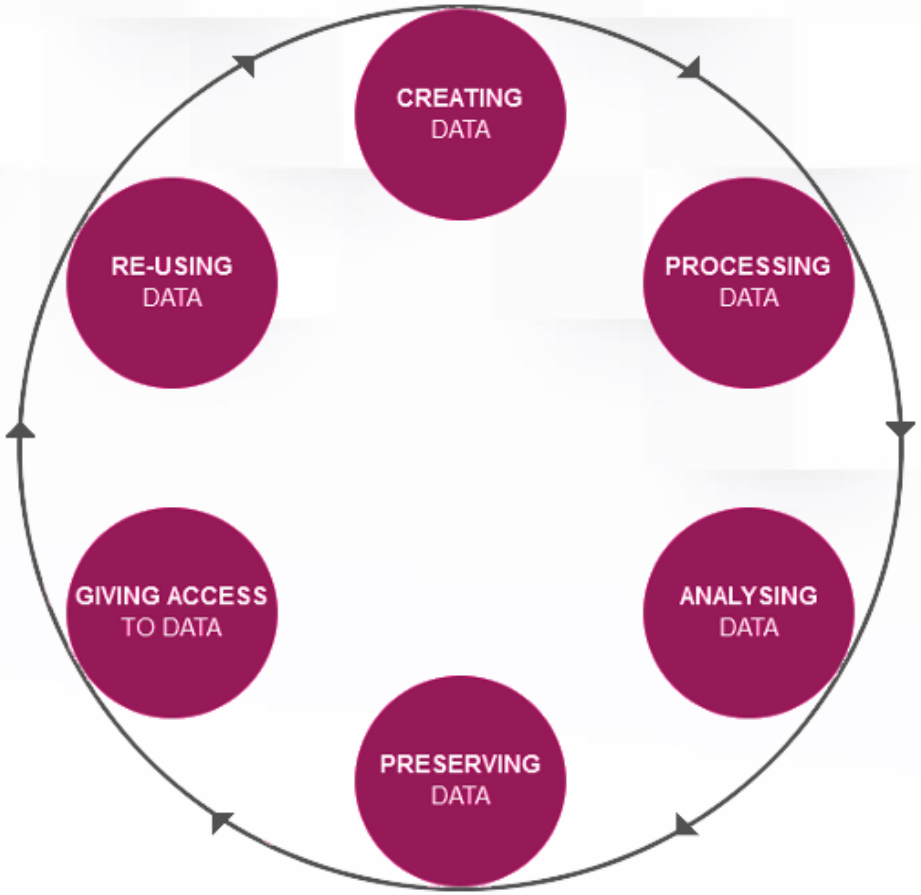
With thanks to John Jungck

Challenge



To go from publications as the single outcome of a research process to…

Challenge



A system where the whole research process is visible and valued.

Why?

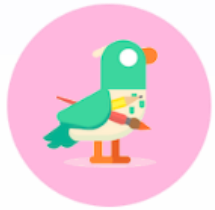
Prevents data loss

Increases transparency

Maximize usefulness

Credit & longer shelf life [1]

Citizens science & collaboration



Write a data paper

1. e.g. Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. PeerJ 1:e175 https://doi.org/10.7717/peerj.175, Piwowar HA, Day RS, Fridsma DB (2007) Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS ONE 2(3): e308. doi:10.1371/journal.pone.0000308

Exempli gratia

Prevents data loss

e.g. “The full profile listings are on floppy disks which are available upon request”*

* Fernolz et al (1989) A survey A survey of measurements and measuring techniques in rapidly distorted compressible turbulent boundary layers.

Maximize usefulness and transparency

Over half of psychology studies fail reproducibility test

Largest replication study to date casts doubt on many published positive results.

Monya Baker

THE ROYAL SOCIETY biology letters

A survey of publication bias within evolutionary ecology

Negative results are disappearing from most disciplines and countries

Daniele Fanelli

OPEN ACCESS Freely available online PLoS one

How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data

Daniele Fanelli*

Most Scientific Research Data From the 1990s Is Lost Forever

A new study has found that as much as 80 percent of the raw scientific data collected by researchers in the early 1990s is gone forever, mostly because no one knows where to find it.

DANIELLE WIENER-BRONNER DECEMBER 23, 2013

Research integrity

"It was a mistake in a spreadsheet that could have been easily overlooked: a few rows left out of an equation to average the values in a column. The spreadsheet was used to draw the conclusion of an influential 2010 economics paper: that public debt of more than 90% of GDP slows down growth. This conclusion was later cited by the International Monetary Fund and the UK Treasury to justify programmes of austerity that have arguably led to riots, poverty and lost jobs."

The error that could subvert George Osborne's austerity programme

The theories on which the chancellor based his cuts policies have been shown to be based on an embarrassing mistake

▲ George Osborne says that Ken Rogoff, the man whose economic error has been uncovered, has strongly influenced his thinking. Photograph: Stefan Wermuth/PA

Science News *from research organizations*

Scientists link biodiversity genomics with museum wisdom through new public databa

Genomic observatories metadatab pact of global challenges across life

Date: August 3, 2017

Source: Smithsonian

Summary: A new publicly available datab ples, making it easier for rese and ecological a when samples w different location research commu

Data Sharing Leads to Powerful Tools for Fighting Fire

The Wildland Fire Decision Support System uses geospatial data and predictions of fire spread to inform decisions on wildland fires.

BY ANDREW BAILEY, NANCY VON MEYER AND BEN BUTLER / DECEMBER 3, 2013

High Park Fire wildland fire crews at the Incident Command Post in Fort Collins, Colo., in June 2012.

MICHAEL RIEGER/FEMA

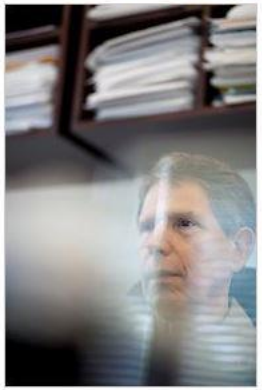
Sharing of Data Leads to Progress on Alzheimer's

By GINA KOLATA
Published: August 12, 2010

In 2003, a group of scientists and executives from the National Institutes of Health, the Food and Drug Administration, the drug and medical-imaging industries, universities and nonprofit groups joined in a project that experts say had no precedent: a collaborative effort to find the biological markers that show the progression of Alzheimer's disease in the human brain.

LINKEDIN
COMMENTS
PRINT
REPRINTS
SHARE

Enlarge This Image



Michael Temchine for The New York Times
Neil Buckholtz, chief of the Dementias of Aging Branch at the National Institute of Aging, in the National Institutes of Health.

Now, the effort is bearing fruit with a wealth of recent scientific papers on the early diagnosis of Alzheimer's using methods like PET scans and tests of spinal fluid. More than 100 studies are under way to test drugs that might slow or stop the disease.

And the collaboration is already serving as a model for similar efforts against Parkinson's disease. A $40 million project to look for biomarkers for Parkinson's, sponsored by the Michael J. Fox Foundation, plans to enroll 600 study subjects in the United States and Europe.

The work on Alzheimer's "is the precedent," said Holly Barkhymer, a spokeswoman for the foundation. "We're really excited."

The key to the Alzheimer's project was an agreement as ambitious as its goal: not just to raise money, not just to do

COVID-19 research

Open access data benefits millions of scientists around the world and is essential for a rapid response to the COVID-19 pandemic

Artist's impression of open access COVID-19 data sharing. Credit: Spencer Phillips, EMBL-EBI.

What is data management?

- Research Data Management is **part of the research process**, and aims to make the research process as efficient as possible

- Research data management concerns **the organisation** of data, from its entry to the research cycle through to the **dissemination** and **archiving** of valuable results.

What is data management?

→ Can people find, access and understand my data?

EXPLAIN IT

STORE IT SAFELY

OPEN IT

→ A Data Management Plan is a tool that can help you think about and address these issues.

What is open data?

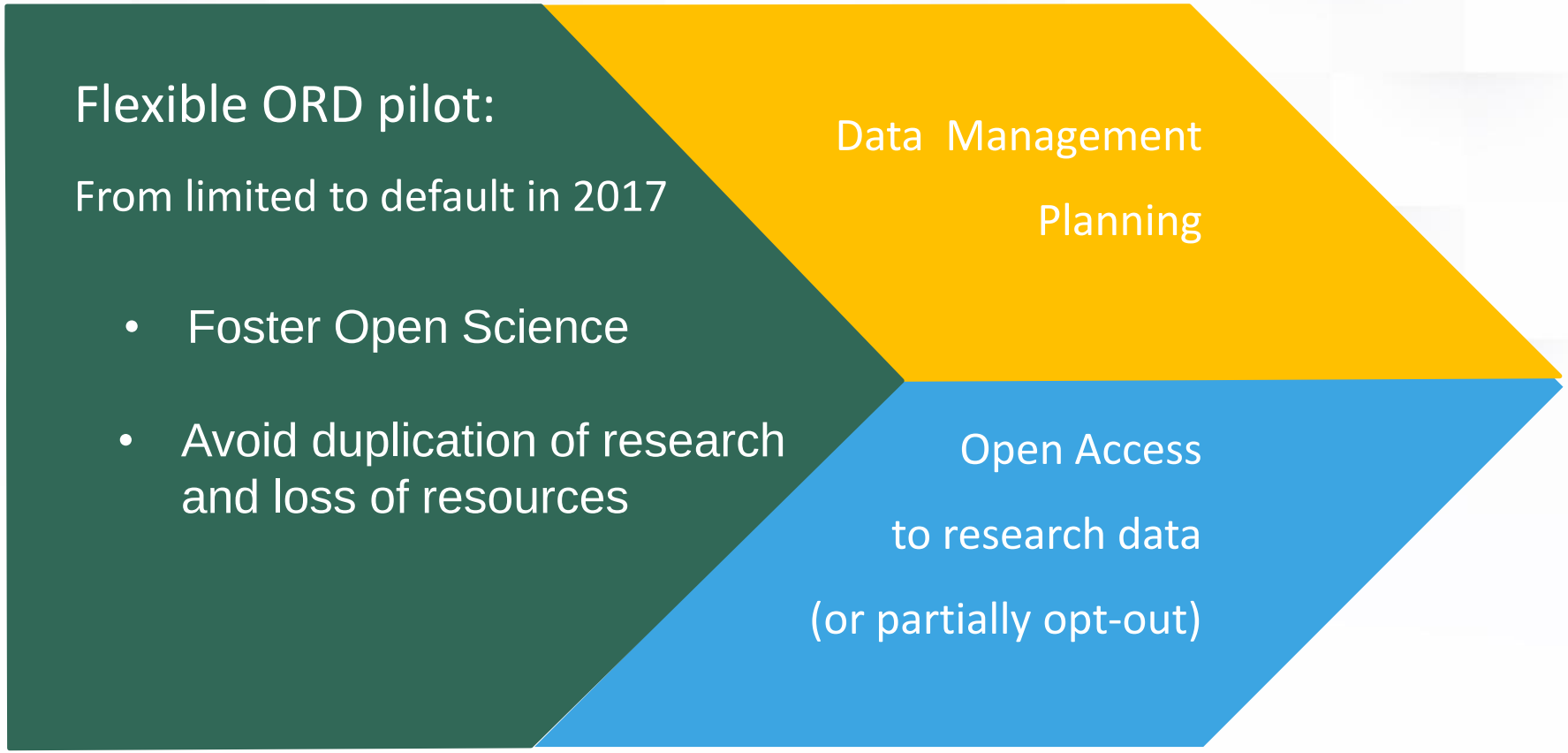
Open Data

Free to access reuse, repurpose, and redistribute

Data sharing

Restricted access to limited amount of people under certain conditions

Horzion 2020 and Open Research Data



Requirements of the Data Pilot

1. Data Management Plan (DMP)

2. Deposit data in data repository

3. Provide information to validate results

4. Open up data

STEP 1

WRITE A DMP

dmponline.be

Update at

- 6 months
- Periodic evaluation
- Final review

STEP 2

FIND REPOSITORY

Matches data needs

Data Repositories

- discipline/institutional
- www.re3data.org
- Zenodo

STEP 3

DEPOSIT DATA

(Open) Data

Metadata

Other tools

- Standard File Formats
- Standards metadata schema
- (Open) Licences

SUPPORT

Supporting infrastructure and information

Open Research Data Pilot

- EC guidelines
- OpenAIRE.eu

FAIR Data Management guidelines

EUROPEAN COMMISSION
Directorate-General for Research & Innovation

H2020 Programme

Guidelines on

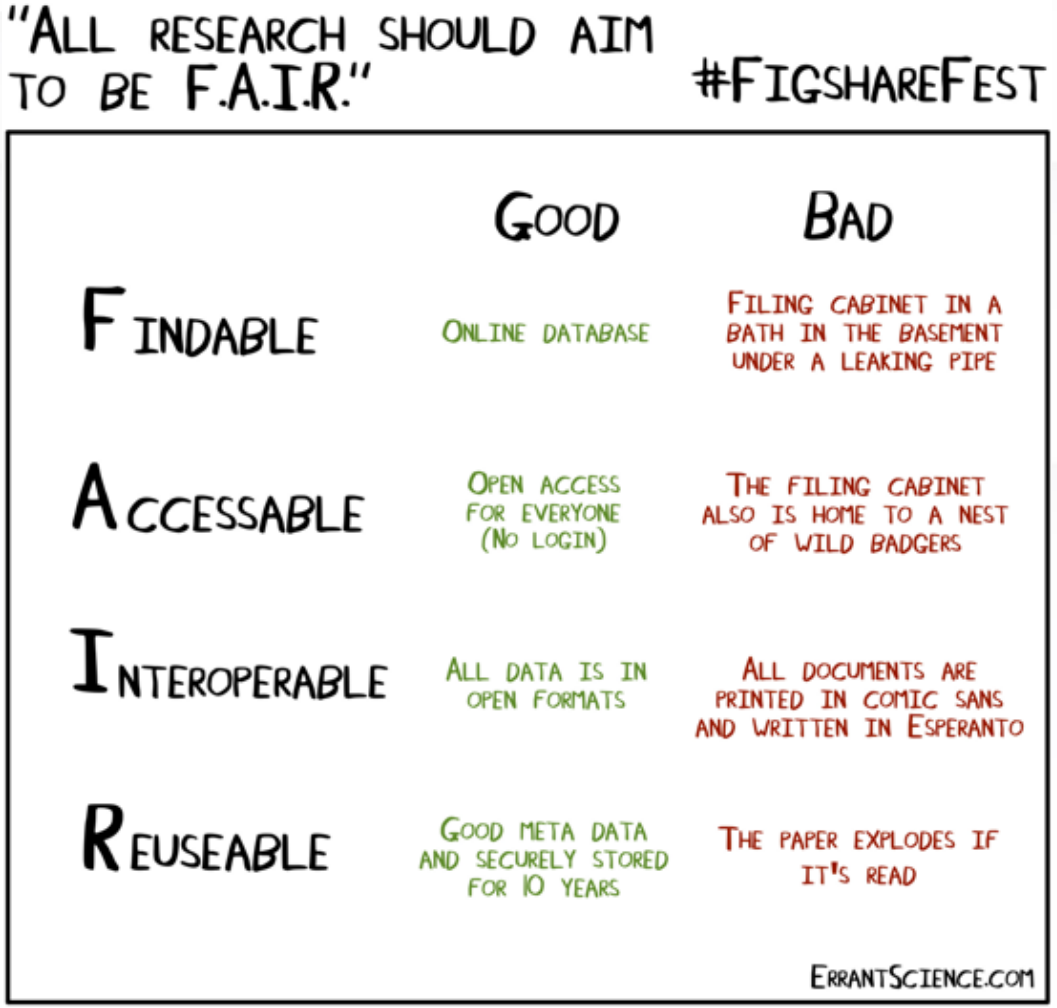
FAIR Data Management in Horizon 2020

Version 3.0
26 July 2016

- Notes the extension of the pilot
- Clarifies concept of FAIR data
- Explains what a DMP is and when they should be updated
- Notes what happens at proposal, submission and evaluation stage
- Explains costs are eligible
- Provides a DMP template

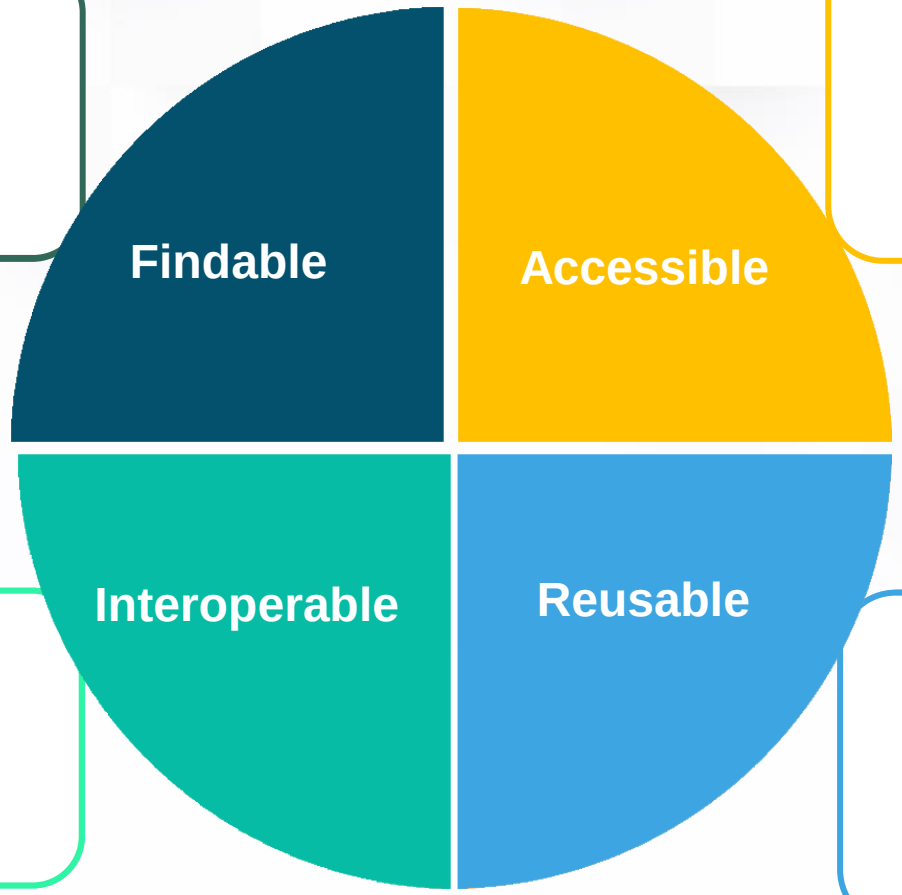
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Guided by a set of principles: FAIR



FAIR data principles

- Metadata
- Persistent identifier
- Naming convention
- Keywords
- Versioning



- Software, documentation
- Data repository

- Standards
- Vocabulary
- Methodologies

- Licensing

Research data in Horizon Europe

Continuation and strengthening of ECs research data mandate with a focus on good Research data management and FAIR data principles.

- Establish a **data management plan** (‘DMP’) and update it
- Deposit the data in a **trusted repository**
- Ensure **open access** following the principle ‘as open as possible as closed as necessary’

Changes from Horizon 2020

- No longer a pilot
- Focus on Research Data Management and **FAIR data**
- **Open (CC-BY or CCO) license**
- Metadata of deposited data must be open under a Creative Common Public Domain Dedication
- ***"EOSC Pilot":***

  → If required in the call conditions, the data repository must be federated in the EOSC in **compliance with EOSC**

EOSC-Pillar

Coordination and Harmonisation of National & Thematic Initiatives to support EOSC

EOSC-Pillar will coordinate national Open Science efforts across Austria, Belgium, France, Germany and Italy, and ensure their contribution and readiness for the implementation of the European Open Science Cloud.

EOSC-Pillar has received funding from the European Union's Horizon 2020 research and innovation Programme under Grant Agreement No. 857650.
This material by the EOSC-Pillar is licensed under a Creative Commons Attribution 4.0 International License.

@EOSCPillar /company/eosc-pillar www.eosc-pillar.eu

The European Open Science Cloud is...

- set up by the **European Commission**
- for **existing scientific data infrastructures**
- it's a **process** (not a project) to build a virtual environment
- to make it easier to **find, access, and reuse** research data, tools and services
- to **interlink** and **combine** data sets
- **across disciplines** and borders
- to **enable new and unexpected discoveries**

EUROPEAN OPEN SCIENCE CLOUD

Challenges EOSC aims to address

- Provide **access to data/results** which are now fragmented
- Handle the **increase in data volume**
- Manage **different data types**
- Facilitate data publications and **sharing of data**
- Foster **multi-disciplinary collaboration**, connecting disciplines that are now working separately

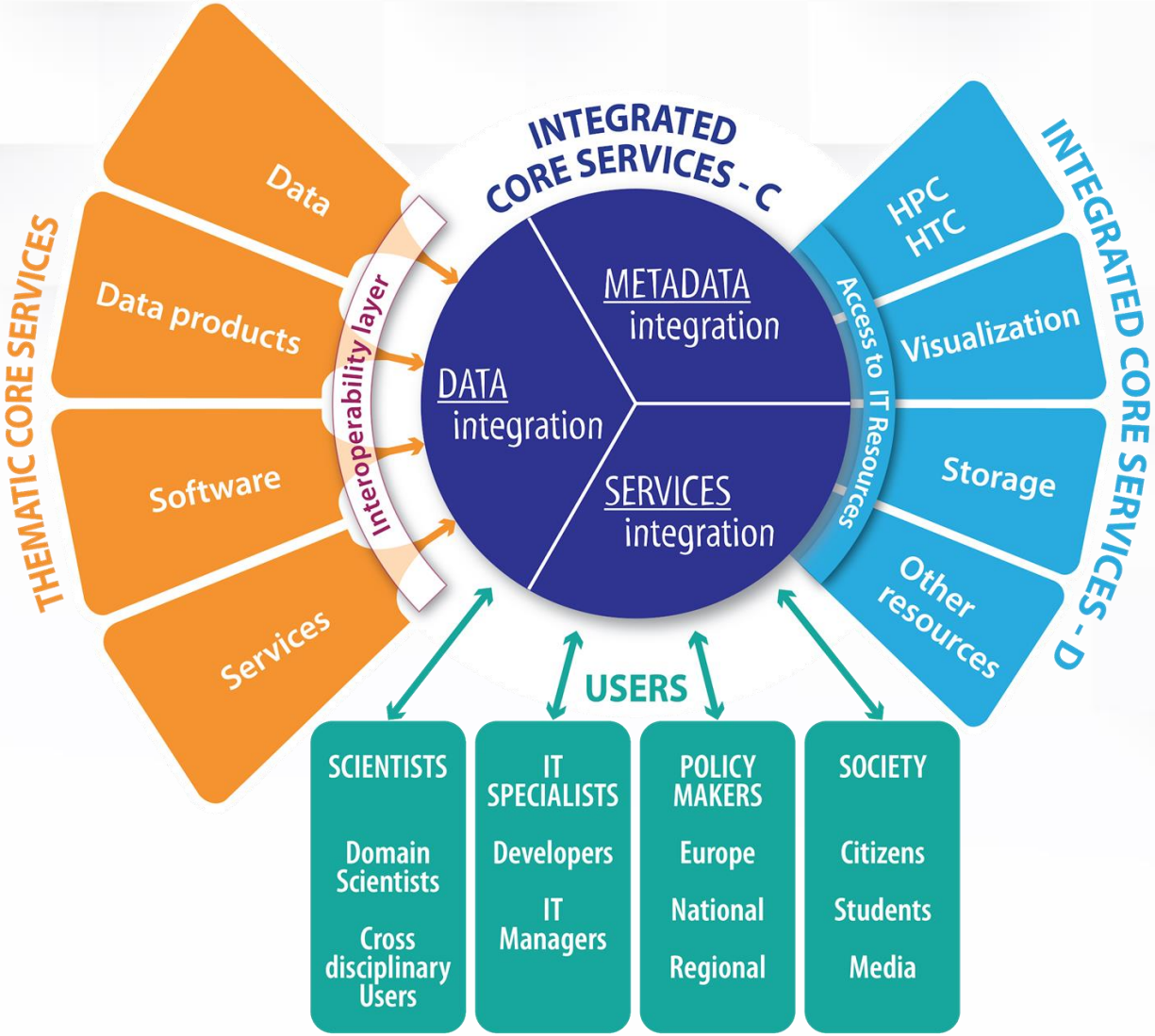
EUROPEAN OPEN SCIENCE CLOUD

"1.7 million European researchers and 70 million professionals in science and technology a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines"

European approach

Stakeholders working together to

- cataloguing datasets
- combining pre-existing research infrastructures
- provide services for researcher data
- exchange datasets across disciplines
- create the internet of FAIR data



Source: https://epos-ip.org/epos-vision-eosc

FAIR is a crucial aspect of the European Open Science Cloud.

FAIR is...

To make the European Open Science Cloud work **FAIR** is not only the **gold standard** for research data BUT also for the infrastructure and services.

EOSC services for researchers

Visit marketplace.eosc-portal.eu

- Data and data **processing**
- Data **storage** and data services
- **Cloud computing**: e.g optimised for genome analysis pipelines
- Containerised **workflows**
- **Training and support** on scientific data management
- **High performance computing** in PRACE, HPC-Europa etc. will become accessible in EOSC as well

Mission

Boost European research by providing **all necessary interfaces and functionalities** to equip EOSC Users with knowledge about and possibility to use the resources available in EOSC for the sake of supporting science across the Europe.

Main functions

Gateway to **EOSC Resources**

- **284** categorised and well described, research supporting, **resources**
- **Generic** resources along with **specialised, thematic** services (both physical and digital)
- Solutions developed for all branches of science
- Resource offers **tailored** to specific use-cases
- 4 types of access mode

Communication channel with EOSC Providers and other EOSC stakeholders

- **160 providers** of different kinds (e-infrastructures, user communities, scientific institutions, etc.)
- **Direct communication** with Providers
- Possibility to request a new EOSC resource compositions or a new resources
- **User feedback** and resource rating
- User support from EOSC Providers and EOSC Experts
- Support for **individual use cases** from EOSC Providers

Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

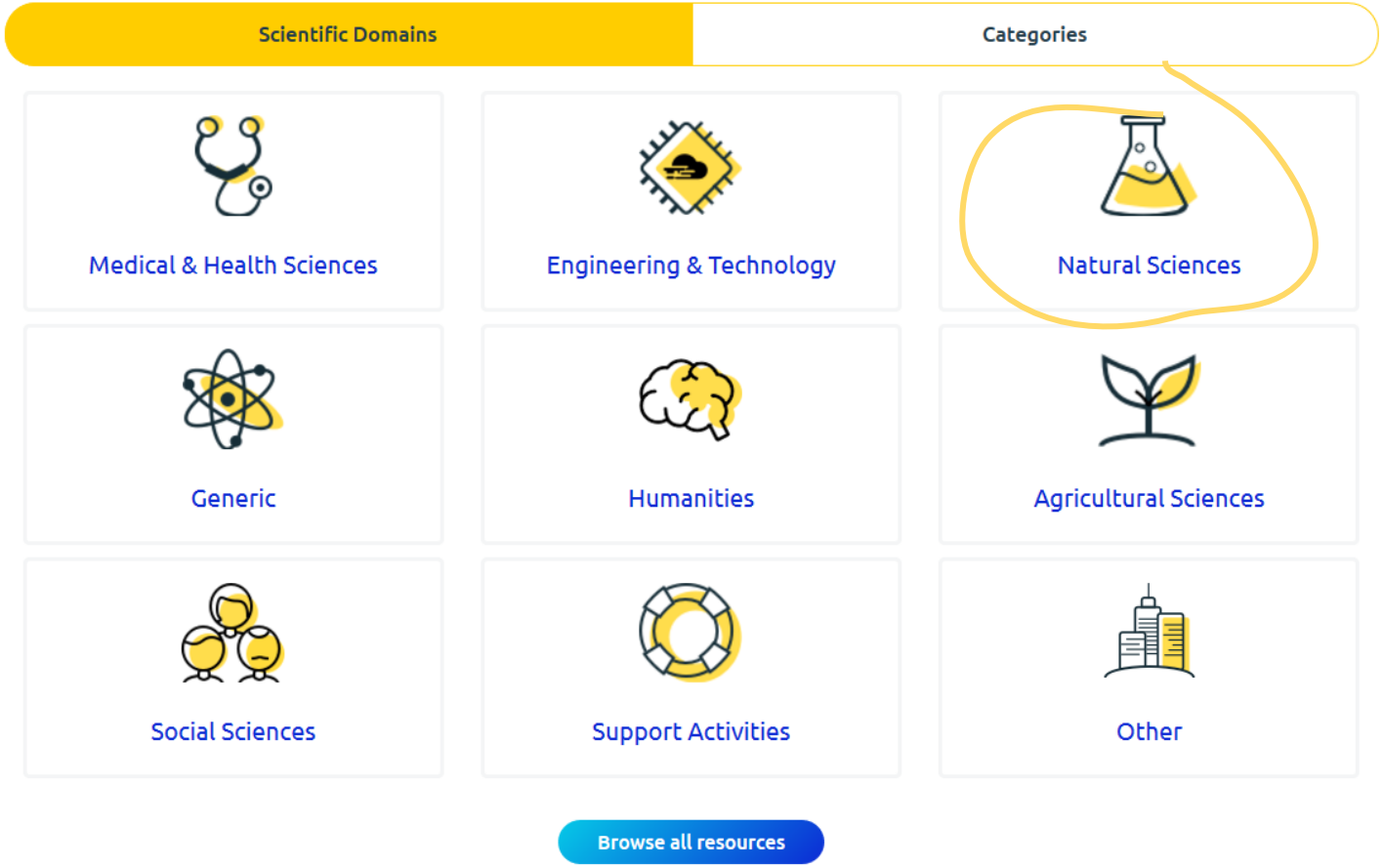
Institutions/researchers will be able to

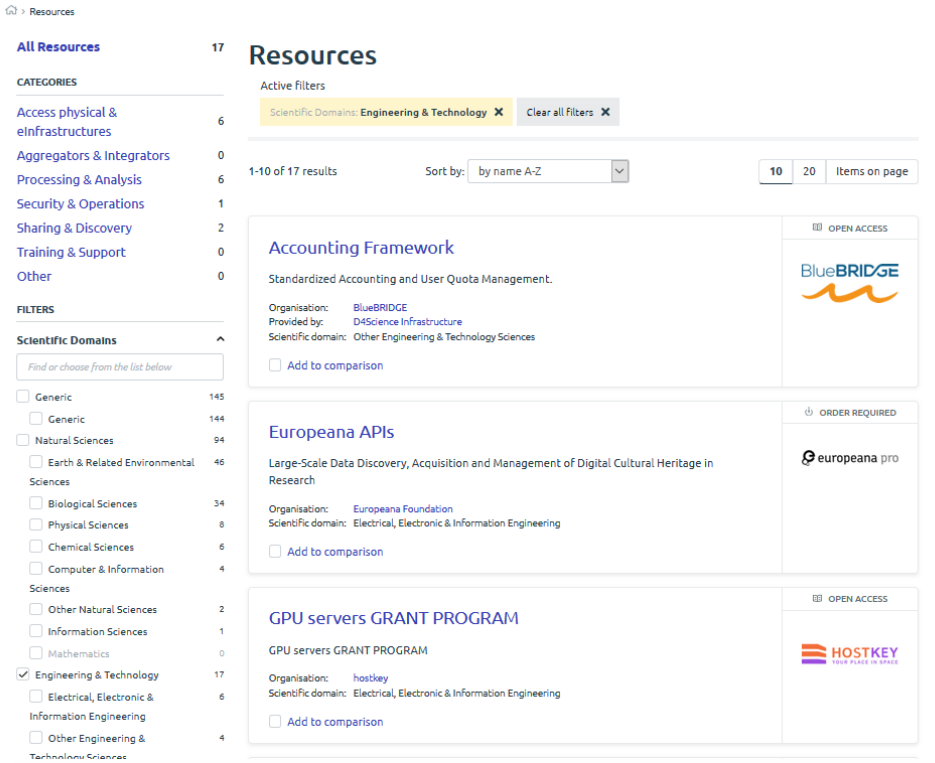
- Find data
- Exchange data
- Have access to data
- Reproduce data
- Interdisciplinary use of data
- Use federated services and infrastructure
  - → knowledge production will also benefit industry and businesses

Demo

https://marketplace.eosc-portal.eu/

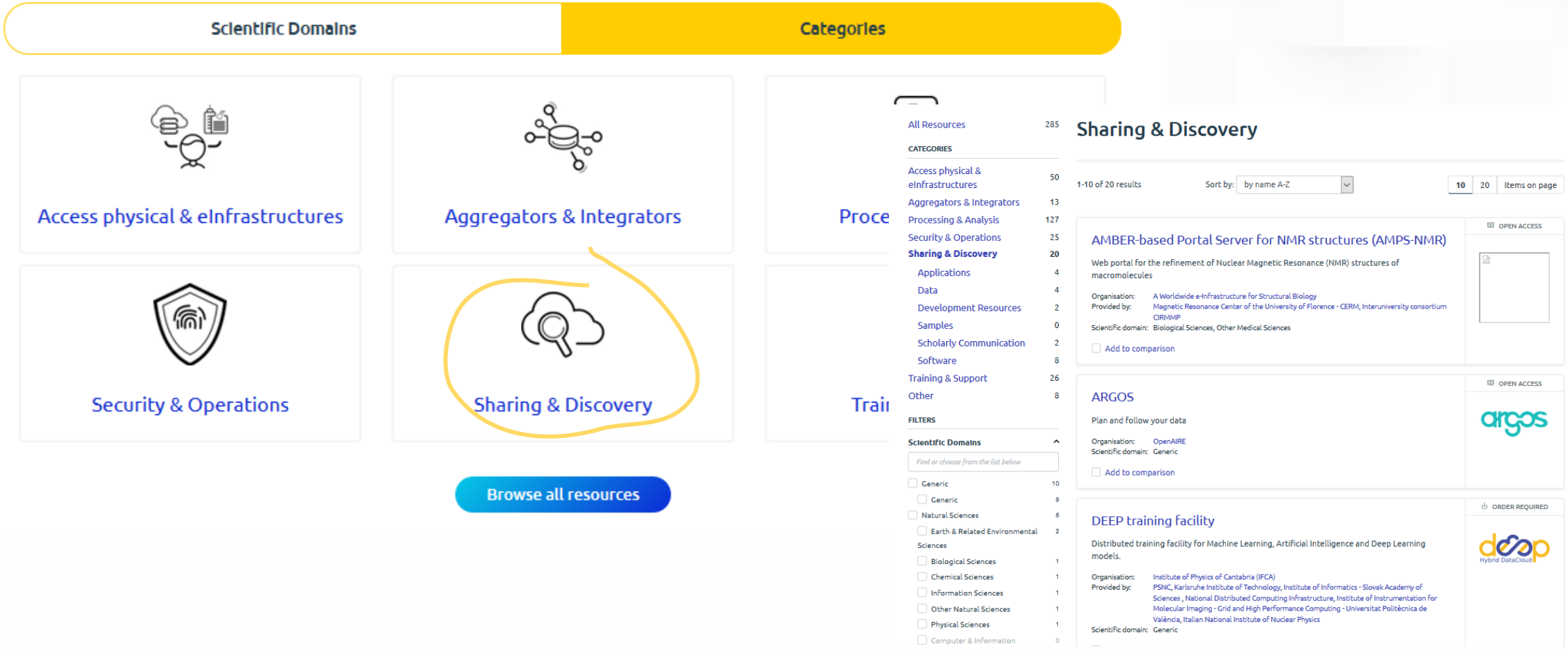
9 scientific domains to start with...





Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

Or 6 categories...



Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

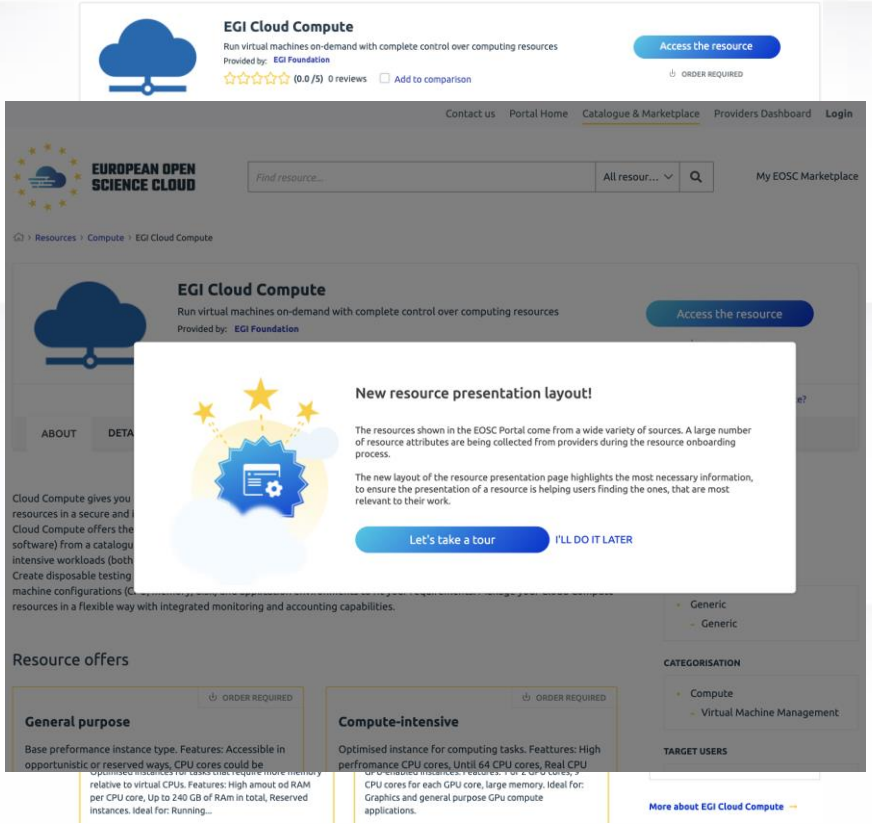
Discovery functionalities

Focus on user perception and information clarity

- Layout designed to promote the data crucial to **assess** the resource

Tour guide to present and explain new layout

- Resource header
- Resource classification
- About section
- Details
- Reviews
- Resource offers



Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

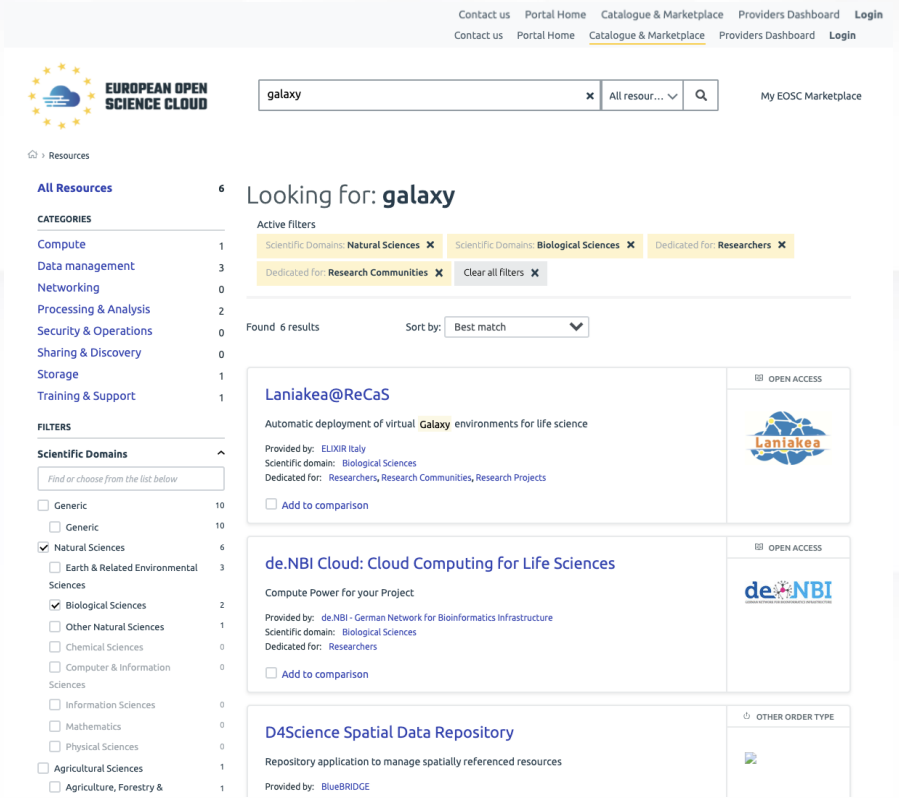
Discovery functionalities

Resource comparison

- **Side-by-side review** of the key resource metadata
- Easy way to **find differences**

Search and filtering

- Enhanced full-text-search methods to find phrases in various resource and offer metadata such as title, tagline, provider fields or offer descriptions
- Resource metadata can be compared by adding resource entries to the comparison tray
- New filters allowing more accurate search



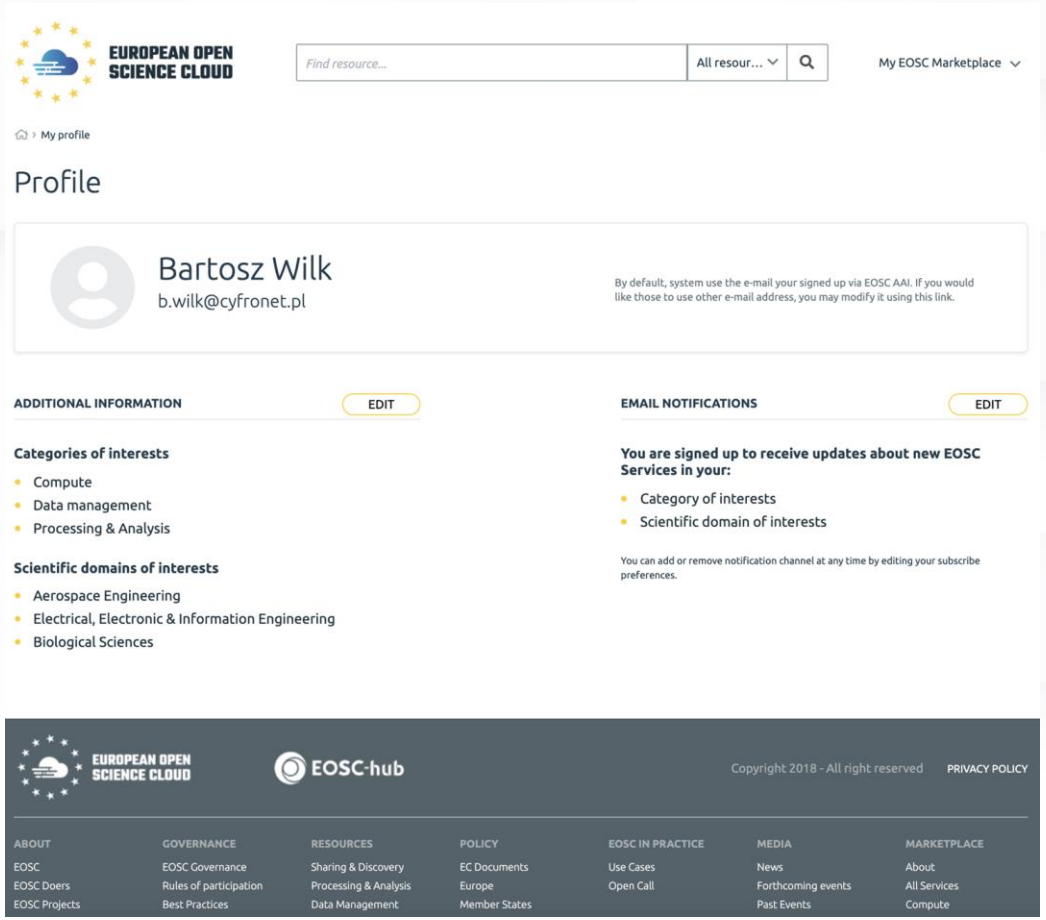
Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

User Profile

- Interface meant to collect **user-relevant data** to serve EOSC user experience-enhancing functionalities
- Subscription for **notifications** about new resources of interest

Coming soon

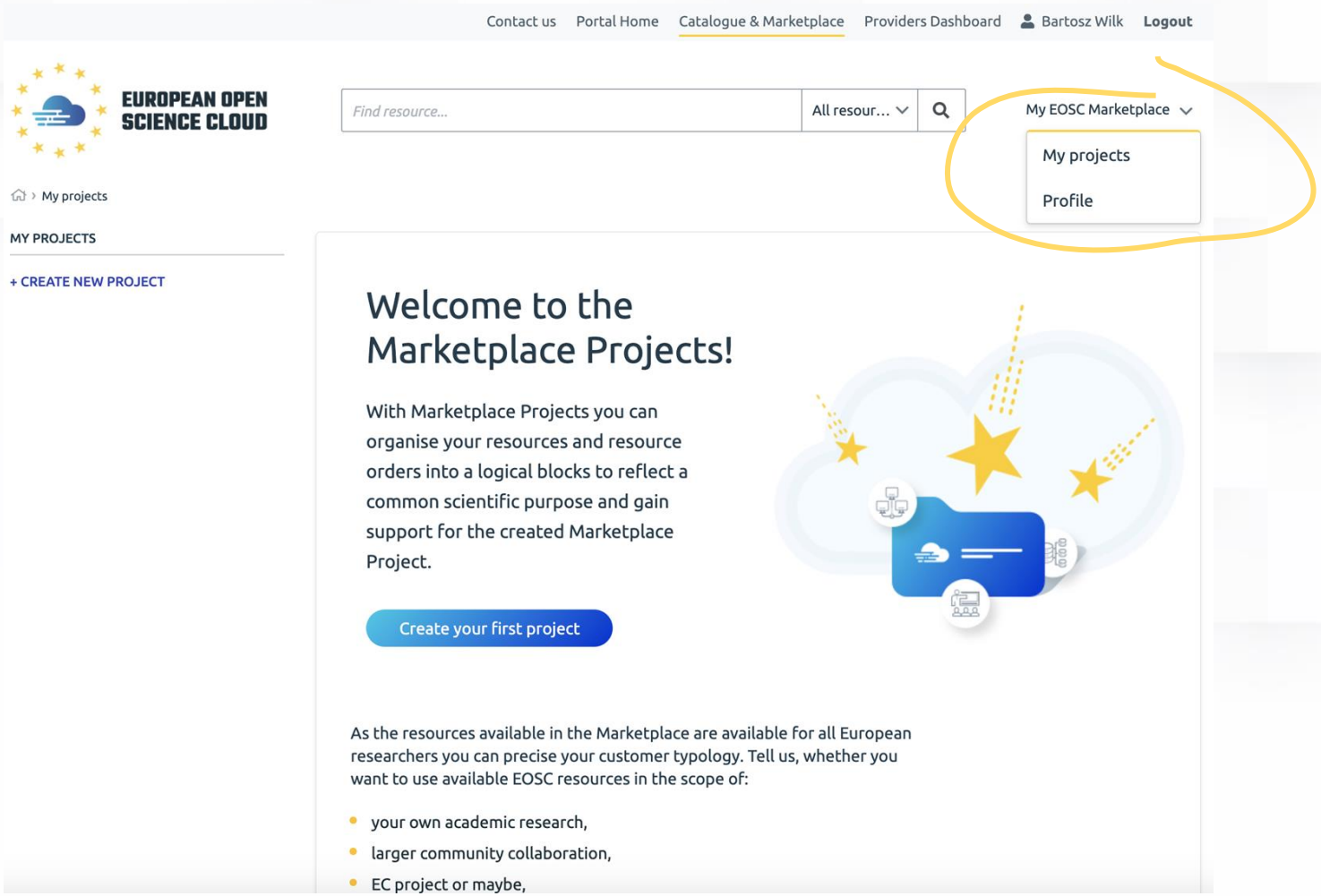
- More user profile properties
- Profile based portal personalization
- Portal recommendations to help in browsing resources



Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

EOSC Marketplace Projects

- **Organise EOSC resources** in a manner that suits your work best
- **Collect resources** for your use case together: easily access the collection of your preferable resources
- **Monitor ordering status** for your EOSC resources
- Interact with the EOSC Experts and EOSC Providers to get **support**
- Rate EOSC resources

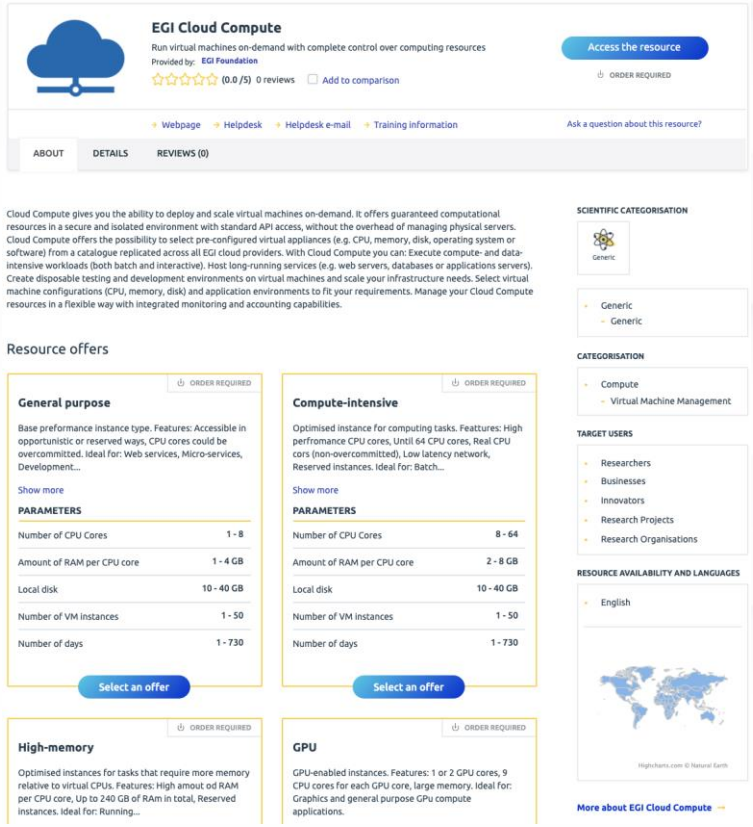


Resource order type

No.	Order Type	What does it mean	MP visibility
1	Request/Order required	Accessing the resource requires undertaking an ordering procedure coordinated by EOSC or the Resource Provider	ORDER REQUIRED
2	Open Access	No ordering procedure necessary to access the resource but user authentication is required	OPEN ACCESS
3	Fully Open Access	Resource can be used without any constraints	FULLY OPEN ACCESS
4	Other	Other/unclassified order type	OTHER ORDER TYPE
5	Various Order Types	Several order types are available for this resource Order type is dependant on the resource offer selection	VARIOUS ORDER TYPES

Accessing the resource

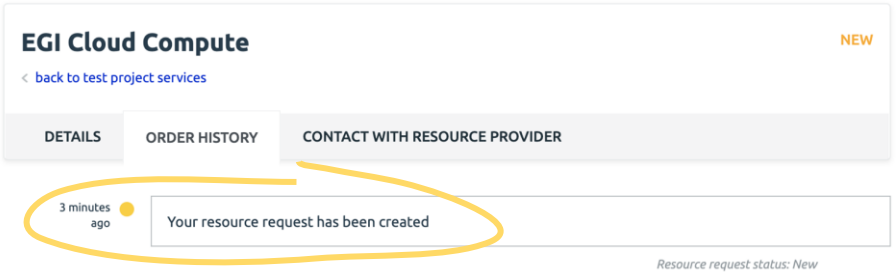
- The main access point for the resource is the **resource presentation page**
- The page allows to analyse resource offering and **choose the offer appropriate** for the desired use case
- Depending on the order type of the resource offer, next steps will be indicated

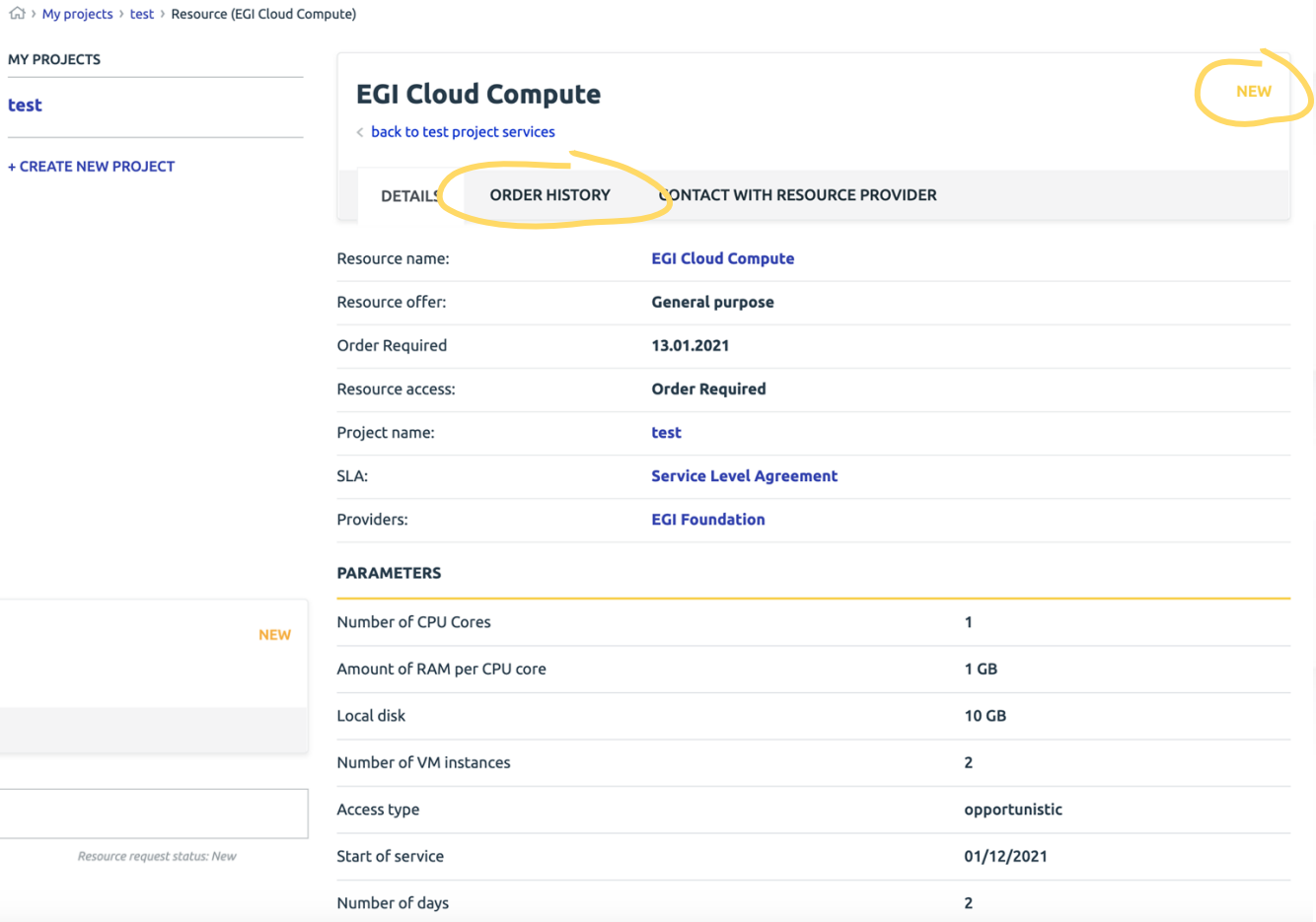


Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

Accessing the resource

- Ordered resources are visible, available to track and use when ready in the project panel
- All resources may be **bookmarked** in the projects for the ease of accessibility





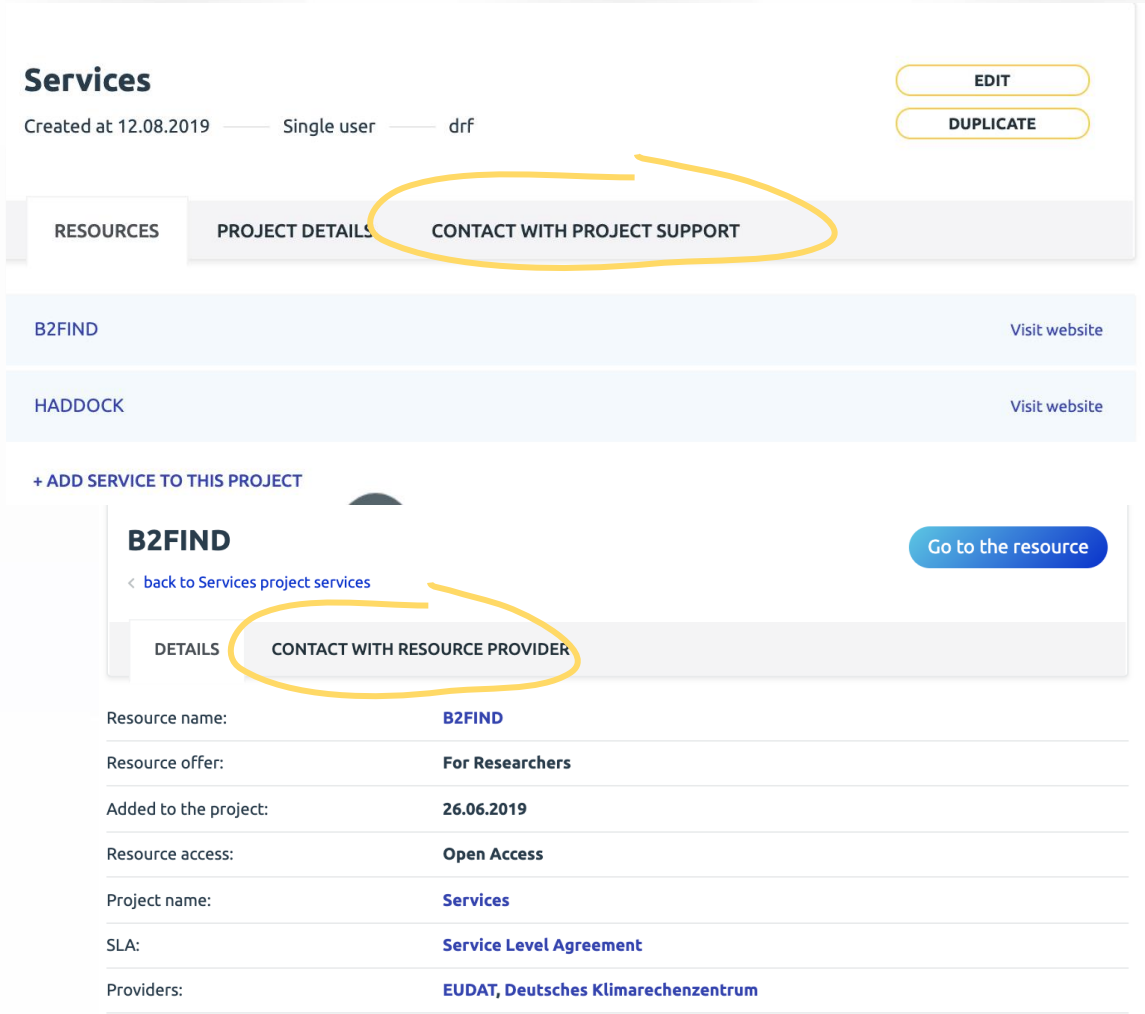
Support channels - projects

Contact with Resource Provider

- Providers support when using the resource
- Incident reporting

Contact with Project Support

- For users looking for **help** in **selecting resources** for their use cases
- Or EOSC experts guidance through **technological or access policy aspects**

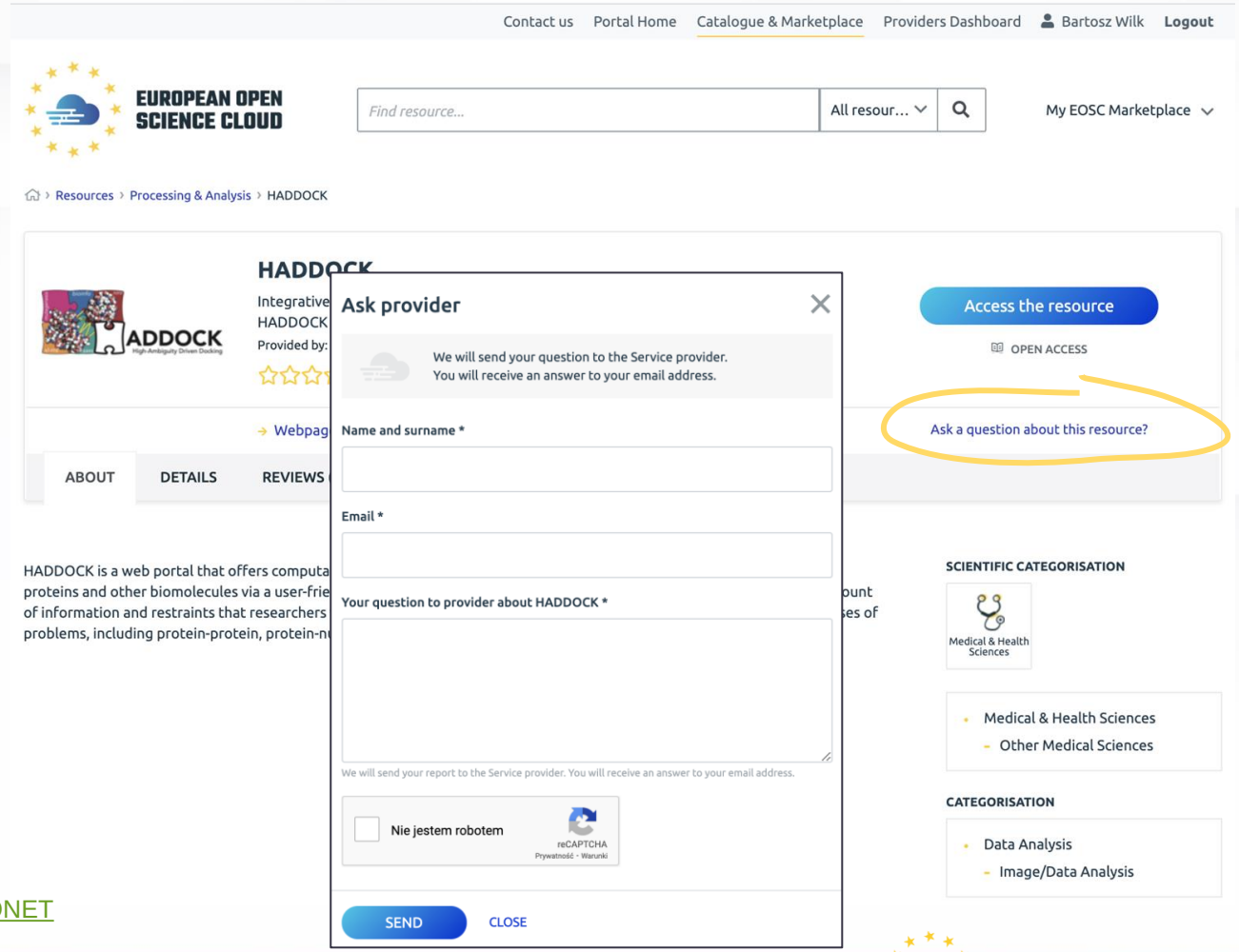


Other support channels

Resource Page - “Ask a Question” button

A user can request clarification or ask a **resource specific question:**

- For authenticated and unauthenticated users
- The question is sent straight to the providers
- The communication is carried out over email

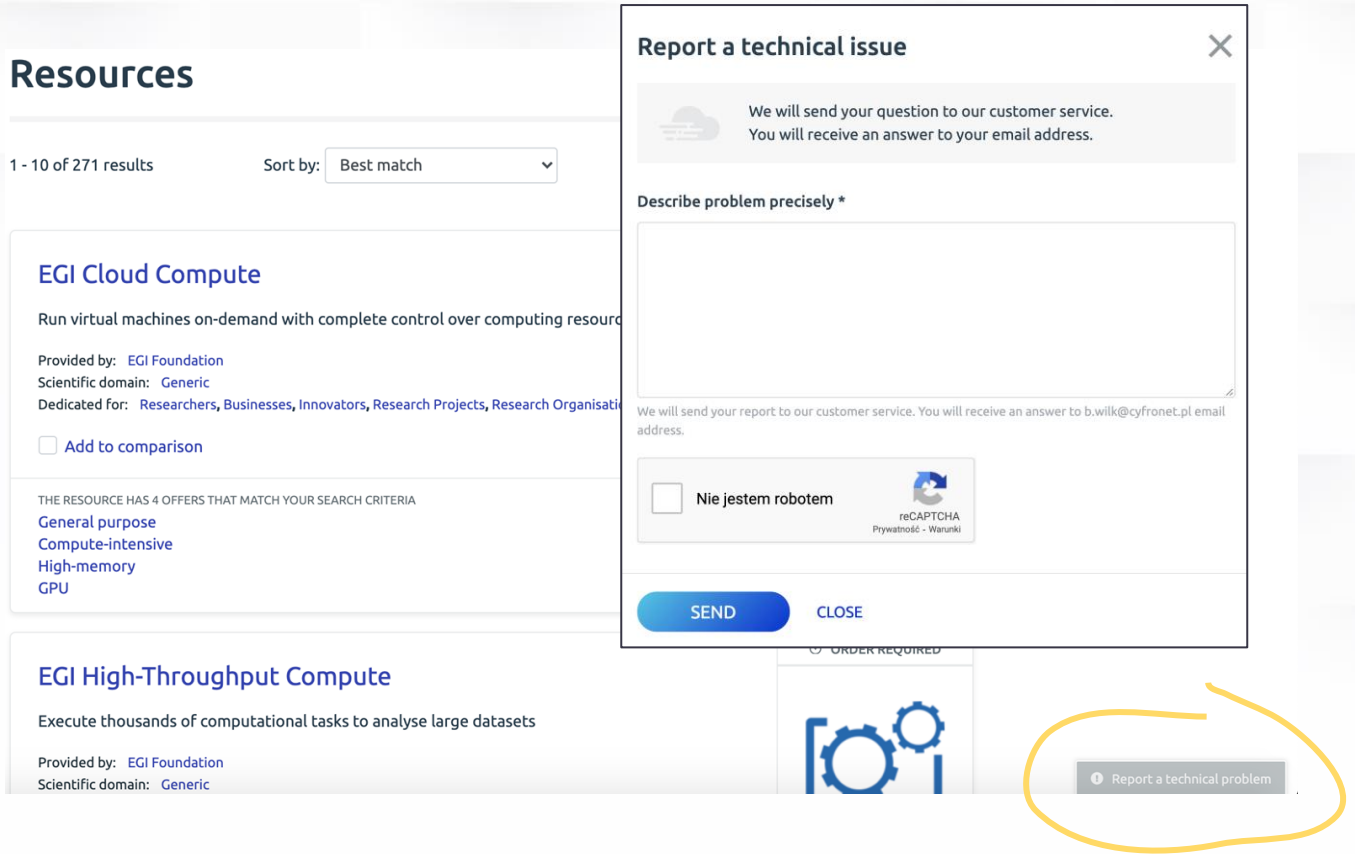


Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

Other support channels

Technical issues

- Report problems with the operation of EOSC Portal: Issues are tracked and handled by the Portal support team

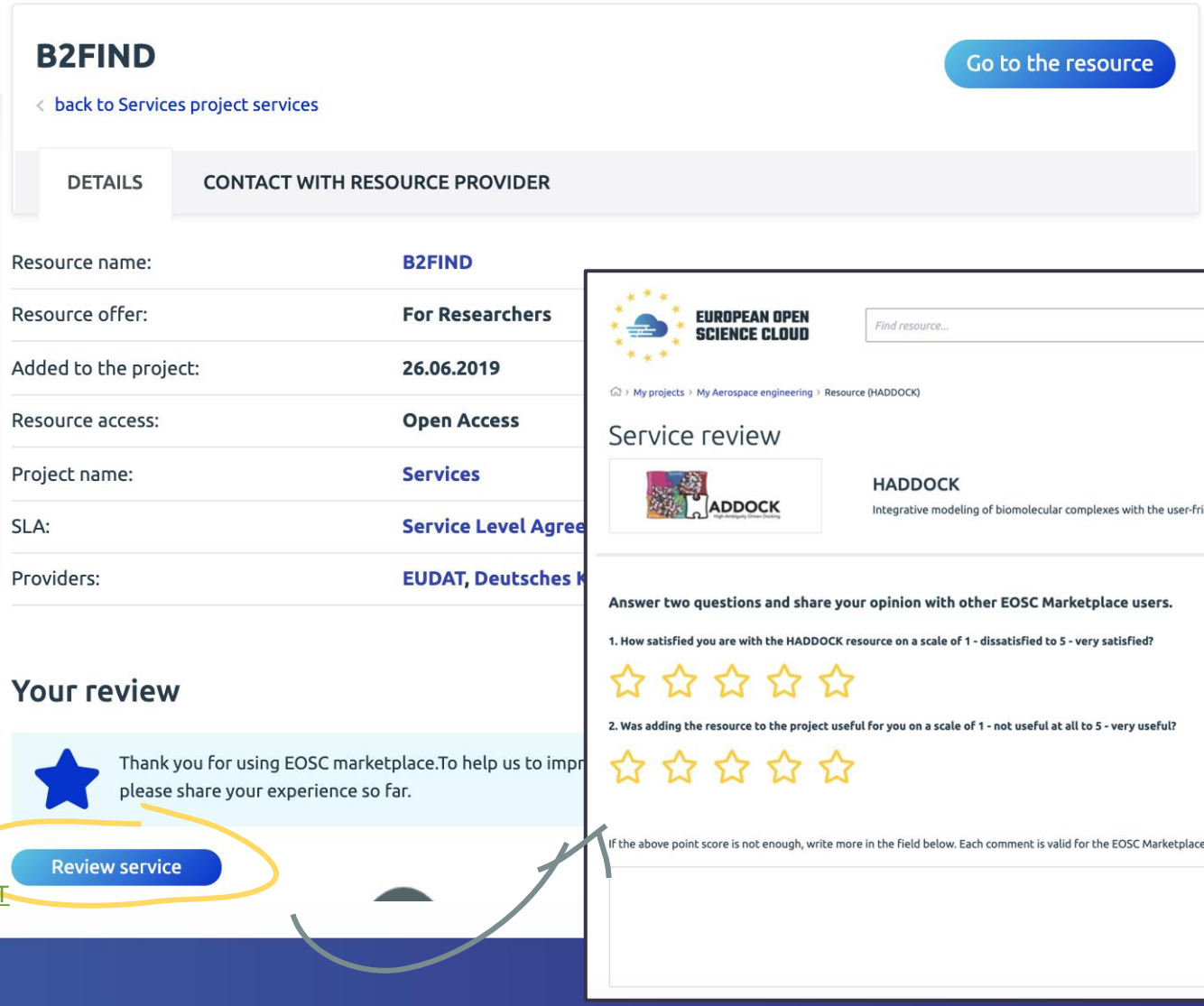


Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

Feedback mechanisms

Service rating and opinions

- Only **used resources** can be rated
- To rate resources -> add them to your Project
- User ratings are used to **measure quality** of the service delivered by the provider

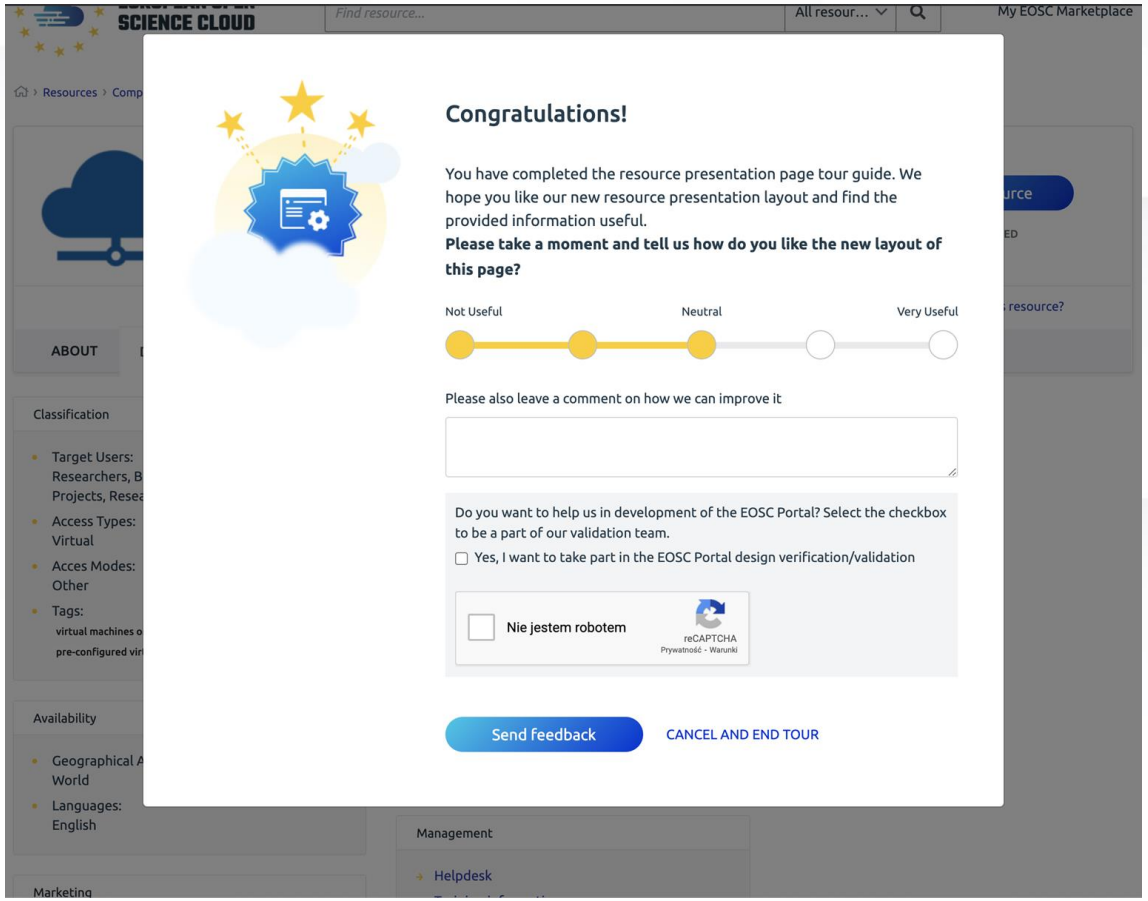


Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

Feedback mechanisms

Features' feedback

- To measure user satisfaction a feedback collection mechanism is employed **at the end of the feature tour guide**
- You can **rate features** and **suggest changes** for the Portal developers
- Create the portal with us!

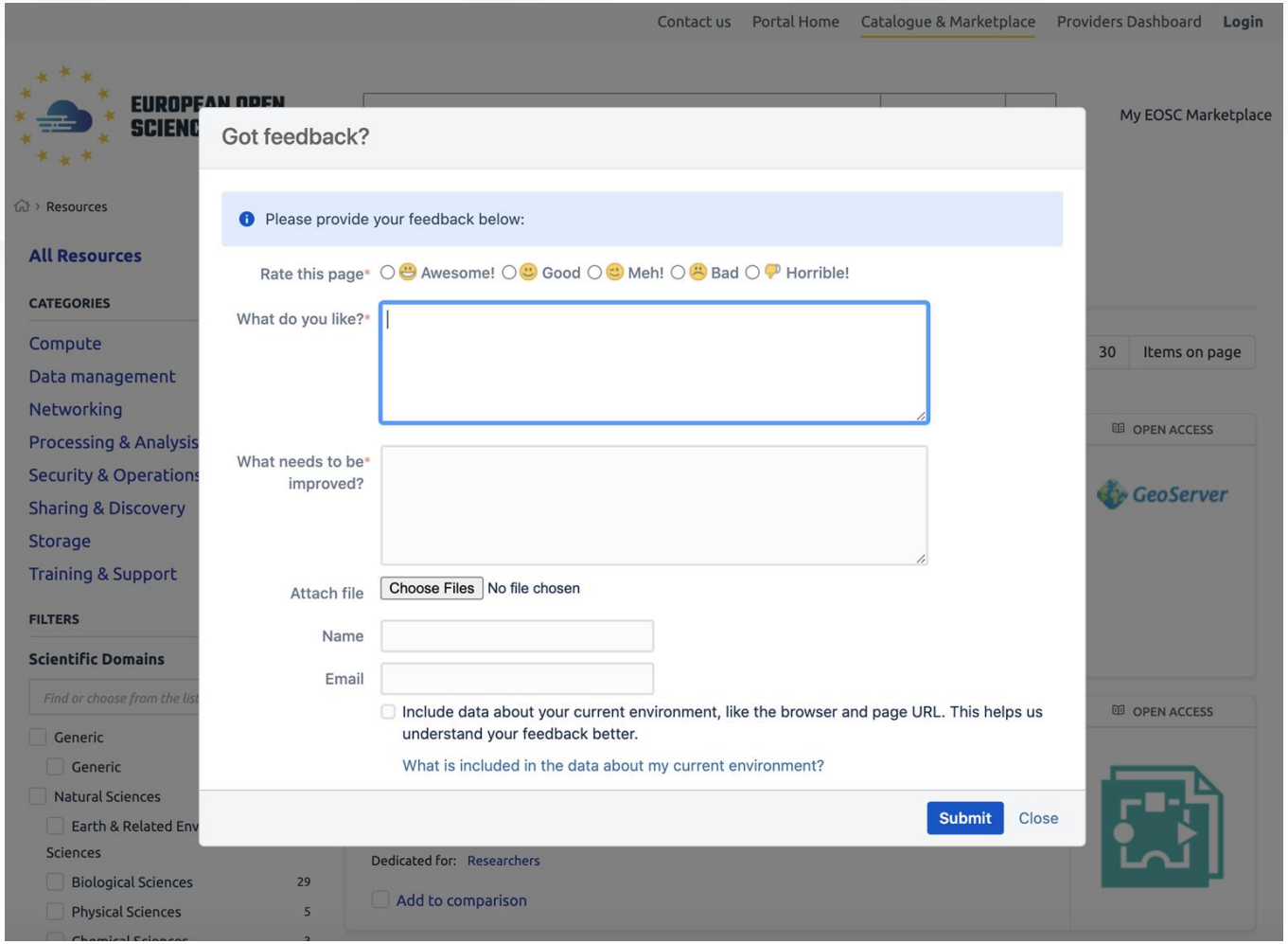


Slides: EOSC-Enhance: EOSC Portal: New user functionalities - **Bartosz Wilk** CYFRONET

Feedback mechanisms

General EOSC Portal feedback mechanism

- “Provide feedback” ribbon
- If you have **any suggestions** for the EOSC Portal development, operational procedures, resource catalogue quality or EOSC in general you can express them directly in the Portal



EOSC-Pillar

Coordination and Harmonisation of National &Thematic Initiatives to support EOSC

Get in touch with us!

www.eosc-pillar.eu

@EoscPillar

/company/eosc-pillar

EOSC-Pillar has received funding from the European Union's Horizon 2020 research and innovation Programme under Grant Agreement No. 857650.