

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

EDSC-PEIDER 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

Open Science Ecosystem





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



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











Increases transparency



Maximize usefulness



Credit & longer shelf life ¹



Citizens science & collaboration



Write a data paper

 e.g. Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. PeerJ 1:e175 <u>https://doi.org/10.7717/peerj.175</u>, 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



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.



Over half of psychology studies fail reproducibility test

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

Monya Baker

THE ROYAL biology SOCIETY

A survey of publication bias within evolutionary ecology

Negative results are disappearing from most disciplines and countries

Daniele Fanelli

OPEN OACCESS Freely available online

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

Maximize usefulness and transparency

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

PLoS one



Daniele Fanelli*



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

Scientists link biodiversity genomics with museum wisdom through new public databa

Genomic observatories metadataba pact of global challenges across life

Date: August 3, 2017 Source: Smithsonian Summary: A new publicly available datab ples, making it essier for record and ecological a when samples w different location research commu

BATTLING EBOLA IN SIERRA LEONE data sharing to improve crisis response

^a Data Sharing Leads to Powerful Tools for Fighting ^w Fire

mmu 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 <u>National</u> <u>Institutes of Health</u>, the <u>Food and Drug Administration</u>, 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 <u>Alzheimer's</u> <u>disease</u> in the human brain.

in	LINKEDIN
ę	COMMENTS
合	PRINT
ē	REPRINTS
+	SHARE
(±	SHARE



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 <u>Parkinson's disease</u>. A \$40 million project to look for biomarkers for Parkinson's, sponsored by the <u>Michael J. Fox Foundation</u>, 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 research on a wast coale, but also to share all the data

THE GLOBAL IMPACT OF OPEN DATA



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?

 \rightarrow Can people find, access and understand my data?

EXPLAIN IT STORE IT SAFELY OPEN IT

 \rightarrow 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

Flexible ORD pilot:

From limited to default in 2017

- Foster Open Science
- Avoid duplication of research and loss of resources

Data Management

Planning

Open Access to research data

(or partially opt-out)





1. Data Management Plan (DMP)



3. Provide information to validate results

2. Deposit data in data repository



4. Open up data





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

EDSC-Pillar Guided by a set of principles: FAIR

"ALL RESEARCH SHOULD AIM **#FIGSHAREFEST** TO BE F.A.I.R." Bad GOOD FINDABLE FILING CABINET IN A ONLINE DATABASE BATH IN THE BASEMENT UNDER A LEAKING PIPE ACCESSABLE OPEN ACCESS THE FILING CABINET FOR EVERYONE ALSO IS HOME TO A NEST (NO LOGIN) OF WILD BADGERS INTEROPERABLE ALL DATA IS IN ALL DOCUMENTS ARE OPEN FORMATS PRINTED IN COMIC SANS AND WRITTEN IN ESPERANTO REUSEABLE GOOD META DATA THE PAPER EXPLODES IF AND SECURELY STORED IT'S READ FOR 10 YEARS ERRANTSCIENCE.COM

EOSC-Pillar

FAIR data principles



EDSC-Pillar 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'

EDSC-Pillar 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":

 \rightarrow If required in the call conditions, the data repository must be federated in the EOSC in compliance with 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.







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



Challenges EOSC aims to address

- Provide access to data/results which are now fragmented
- Handle the increase in data volume
- Manage different data types

EOSC-Pillar

- Facilitate data publications and sharing of data
- Foster **multi-disciplinary collaboration**, connecting disciplines that are now working separately





"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"

LINES





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







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

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



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





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



9 scientific domains to start with...



	17	Resources Active filters		
l & 25	6	Scientific Domain: Engineering & Technology X Clear all filters X		
Integrators nalysis rations	0 6 1	1-10 of 17 results Sort by: by name A-Z	10 20 Items on page	
overy 2 port 0		Accounting Framework	OPEN ACCESS	
	0	Standardized Accounting and User Quota Management. Organisation: BlueBRIDCE Provided by: D4ScienceInfrastructure		
Ins	^	Scientific domain: Other Engineering & Technology Sciences		
m the list below		Add to comparison		
	145			
	144		U ORDER REQUIRED	
es	94	Europeana APIs		
ated Environmental	46	Large-Scale Data Discovery, Acquisition and Management of Digital Cultural Heritage in Research	& europeana pro	
ciences	34	Organisation: Europeana Foundation		
ences	8	Scientific domain: Electrical, Electronic & Information Engineering		
tiences	6	Add to comparison		
k Information	4			
			OPEN ACCESS	
ral Sciences	2	GPU servers GRANT PROGRAM		
Sciences	1			
25	0	GPU servers GRANT PROGRAM	HOSTKEY	
Technology	17	Organisation: hostkey		
lectronic &	6	Scientific domain: Electrical, Electronic & Information Engineering		
gineering		Add to comparison		
neering &	4			



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



Or 6 categories...

Scientific Domains		Categories
Access physical & eInfrastructures	Aggregators & Integrators	All Resources 285 CATECORIES Access physical & 50 eInfrastructures 50 Aggregators & Integrators 13 Processing & Analysis 127 U OPEN Access
		Security & Operations 25 AMBER-based Portal Server for NMR structures (AMPS-NMR) Sharing & Discovery 20 Web portal for the refinement of Nuclear Magnetic Resonance (NMR) structures of macromolecules Applications 4 Organisation: A Worldwide e-Infrastructure for Structural Biology Data 4 Organisation: A Worldwide e-Infrastructure for Structural Biology Development Resources 2 Provided by: Magnetic Resonance Center of the University of Florence - CERM, Interuniversity consortium Samples 0 Scientific domain: Biological Sciences, Other Medical Sciences Scholarly Communication 2 Add to comparison Software 8 Imagenetic Resonance
Security & Operations	Sharing & Discovery	Training & Support 26 Other 8 Plan and follow your data Scientific Domains Find or choose from the list below Generic Ceneric
	Browse all resources	Chenicit P Generic P Natural Sciences P Distributed training facility Distributed training facility for Machine Learning, Artificial Intelligence and Deep Learning models. Distributed sciences 1 Chemical Sciences 1 Organisation: Institute of Physics of Cartabria (IFCA) Provided by: PSNC, Karturule Institute of Technology, Institute of Instrumentation for Moleculare, Institute of Instrumentation for Moleculare, Institute of Instrumentation for Moleculare, Institute of Instrumentation for Molecular Highing Fordinal High Performance Computing Universitat Politiencia de Valencia, Italian National Institute of Nuclear Physics Physical Sciences 1 Other Natural Sciences 1 Physical Sciences 1 Other Natural Sciences 1 Physical Sciences 1 Scientific domain: Centerie Scientific domain: Centerie

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






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







User Profile

- Interface meant to collect userrelevant 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

EUROPEAN OPEN SCIENCE CLOUD	All resour V Q	My EOSC Marketplace 🗸
Ĝ > My profile		
Profile		
Bartosz Wilk b.wilk@cyfronet.pl	By default, system use the e-mail your signed up via E like those to use other e-mail address, you may modif	
ADDITIONAL INFORMATION	EMAIL NOTIFICATIONS	EDIT
	You are signed up to receive updates	
Categories of interests Compute	You are signed up to receive updates Services in your:	
Categories of interests Compute Data management	You are signed up to receive updates	
Categories of interests Compute Data management Processing & Analysis Scientific domains of interests	You are signed up to receive updates Services in your: Category of interests	about new EOSC
Categories of interests Compute Data management Processing & Analysis	You are signed up to receive updates Services in your: Category of interests Scientific domain of interests You can add or remove notification channel at any time	about new EOSC

	DPEAN OPEN NCE CLOUD	DEOSC-hub			Copyright 2018 - All right r	eserved PRIVACY POLICY
	EOSC Governance	Sharing & Discovery	EC Documents	Use Cases	News	About
EOSC Doers	Rules of participation	Processing & Analysis	Europe	Open Call	Forthcoming events	All Services
EOSC Projects	Best Practices	Data Management	Member States		Past Events	Compute





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



MY PROJECTS

+ CREATE NEW PROJECT



Contact us Portal Home Catalogue & Marketplace Providers Dashboard & Bartosz Wilk Logout

- larger community collaboration,
- EC project or maybe,





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	
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	III 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







Accessing the resource

EGI Cloud Compute

- 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

EGI Cloud	-		
DETAILS	ORDER HISTORY	CONTACT WITH RESOURCE PROVIDER	
3 minutes ago	Your resource req	uest has been created	
			Resource request statu

↔ My projects > test > Resource (EGI Cloud Compute)

MY PROJECTS

+ CREATE NEW PROJECT

test

	< back to test project services		
	DETAIL: ORDER HISTORY	CONTACT WITH RESOURCE PROVIDER	
	Resource name:	EGI Cloud Compute	
	Resource offer:	General purpose	
	Order Required	13.01.2021	
	Resource access:	Order Required	
	Project name:	test	
	SLA:	Service Level Agreement	
	Providers:	EGI Foundation	
	PARAMETERS		
NEW	Number of CPU Cores		1
	Amount of RAM per CPU core		1 GB
	Local disk		10 GB
	Number of VM instances		2
	Access type		opportunistic
itus: New	Start of service		01/12/2021
	Number of days		2



NEW



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

Services Created at 12.08.2019 Single user	drf	EDIT DUPLICATE
RESOURCES PROJECT DETAILS	CONTACT WITH PROJECT SUPPORT	
B2FIND		Visit website
HADDOCK		Visit website
+ ADD SERVICE TO THIS PROJECT B2FIND C back to Services project services DETAILS CONTACT WI	TH RESOURCE PROVIDER	Go to the resource
Resource name:	B2FIND	
Resource offer:	For Researchers	
Added to the project:	26.06.2019	
Resource access:	Open Access	
Project name:	Services	
SLA:	Service Level Agreement	
Providers:	EUDAT, Deutsches Klimarechenzentrum	





Other support channels

Contact us Portal Home Catalogue & Marketplace Providers Dashboard 💄 Bartosz Wilk Logout

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





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

Resources	Report a technical issue X
- 10 of 271 results Sort by: Best match 🗸	We will send your question to our customer service. You will receive an answer to your email address.
EGI Cloud Compute Run virtual machines on-demand with complete control over computing resour Provided by: EGI Foundation Scientific domain: Generic Dedicated for: Researchers, Businesses, Innovators, Research Projects, Research Organisat Add to comparison THE RESOURCE HAS 4 OFFERS THAT MATCH YOUR SEARCH CRITERIA General purpose Compute-intensive High-memory GPU	ti We will send your report to our customer service. You will receive an answer to b.wilk@cyfronet.pl email address. Nie jestem robotem recAPTCHA Pryvatność - Warusk
EGI High-Throughput Compute Execute thousands of computational tasks to analyse large datasets Provided by: EGI Foundation Scientific domain: Generic	SEND CLOSE





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

B2FIND	project services			Go to the resource	
DETAILS	CONTACT WITH RESOURCE P	ROVIDER			
Resource name:	B2FIND	,			
Resource offer:	For Res	searchers	EUROPEAN OPEN	Find resource	All resour
Added to the projec	t: 26.06.2		projects > My Aerospace engineering > R		
Resource access:	Open A	cress	vice review	esource (naurock)	
Project name:	Service	es		HADDOCK	
SLA:	Service	e Level Agree		Integrative modeling of biomolecular complexes with the user-frier	ndly, EGI HTC-eni
Providers:	EUDAT,	, Deutsches K		your opinion with other EOSC Marketplace users.	
Your review		1. How s	atisfied you are with the HADDOG	CK resource on a scale of 1 - dissatisfied to 5 - very satisfied?	
	ou for using EOSC marketplace.To h hare your experience so far.	help us to impr 🔥 🏠	☆ ☆ ☆ t	3	
Review service		If the ab	ove point score is not enough, write	more in the field below. Each comment is valid for the EOSC Marketplace o	community.



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!





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

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

* * *		
🎽 👞 📜 EUROPF	AN INDEN	My EOSC Marketplace
SCIENC	Got feedback?	Hy Lobe Harketplace
* * *		
ि⇔ Resources	Please provide your feedback below:	
All Resources	Rate this page* 🔿 😃 Awesome! 🔿 😃 Good 🔿 🥴 Meh! 🔿 🙁 Bad 🔿 ም Horrible!	
CATEGORIES		
Compute	What do you like?*	
Data management		30 Items on page
Networking		
Processing & Analysis		OPEN ACCESS
Security & Operations	What needs to be*	
Sharing & Discovery	improved?	GeoServer
Storage		
Training & Support		
indining a support	Attach file Choose Files No file chosen	
FILTERS	Name	
Scientific Domains		
Find or choose from the list	Email	m
Generic	Include data about your current environment, like the browser and page URL. This helps us understand your feedback better.	OPEN ACCESS
Generic	What is included in the data about my current environment?	
Natural Sciences		
Earth & Related Env	Submit Close	
Sciences	Dedicated for: Researchers	ا نما
Biological Sciences	29 Add to comparison	
Physical Sciences	5 Add to companyon	
Chamical Sciences	2	



EDSC-PILOT Coordination and Harmonisation of National &Thematic Initiatives to support EOSC

Get in touch with us!

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

