

graph.openaire.eu

The OpenAIRE Research Graph

Paolo Manghi

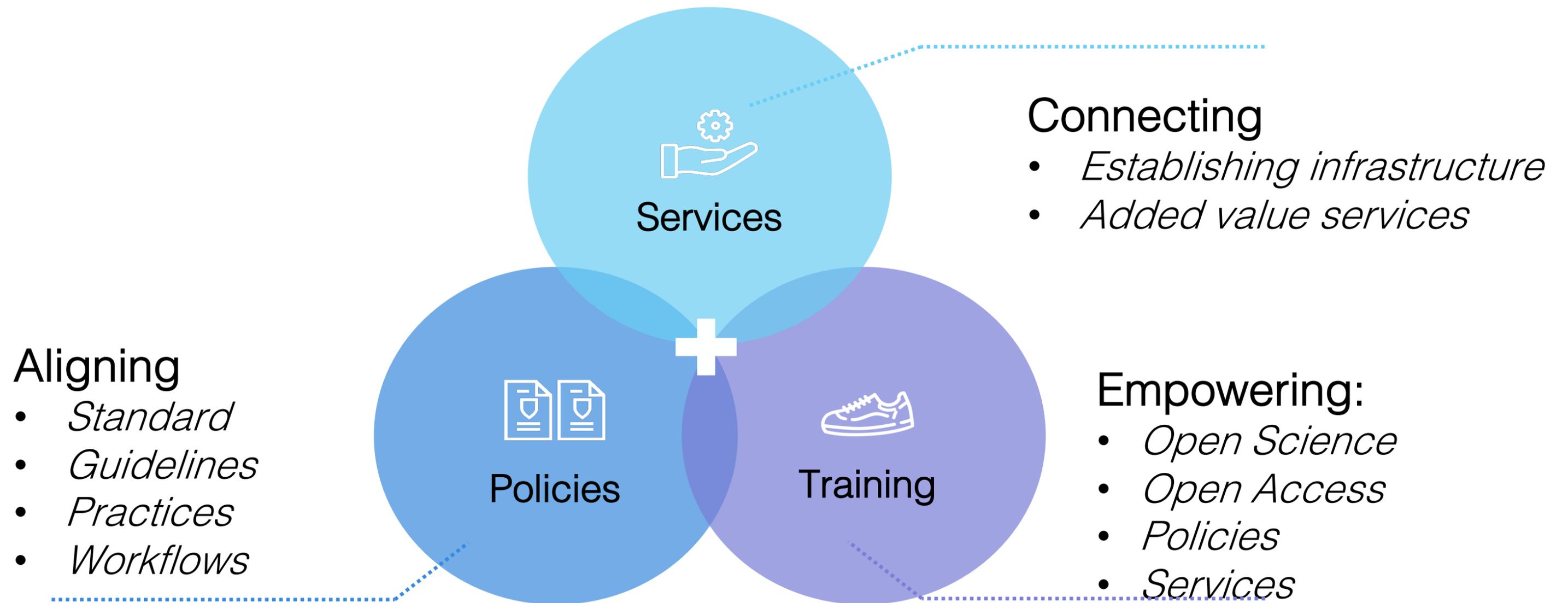
OpenAIRE CTO

Institute of Information Science and Technologies, National Research Council, Italy

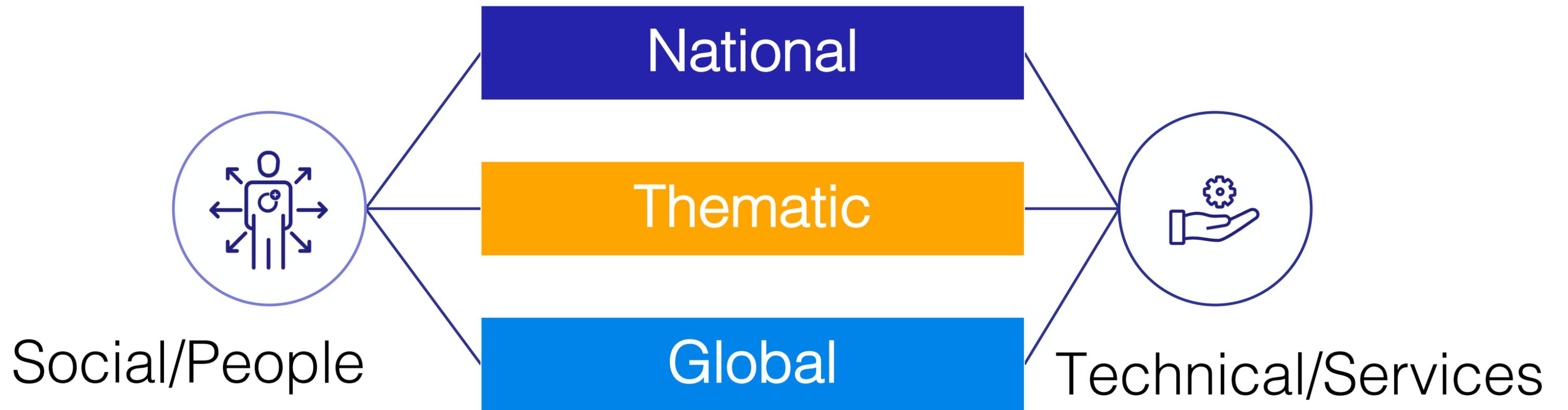


@openaire_
eu

OpenAIRE: three pillars of action



OpenAIRE: three levels of operation



Principles of Open
Science Infrastructures



Tracking Open Science

Reproducibility and transparency require tracking of all outcomes of science and related "context"



Monitoring Open Science

Monitoring quality, impact, and "open scienceness" of science should be a transparent, reproducible process for all, inclusive of research "context"

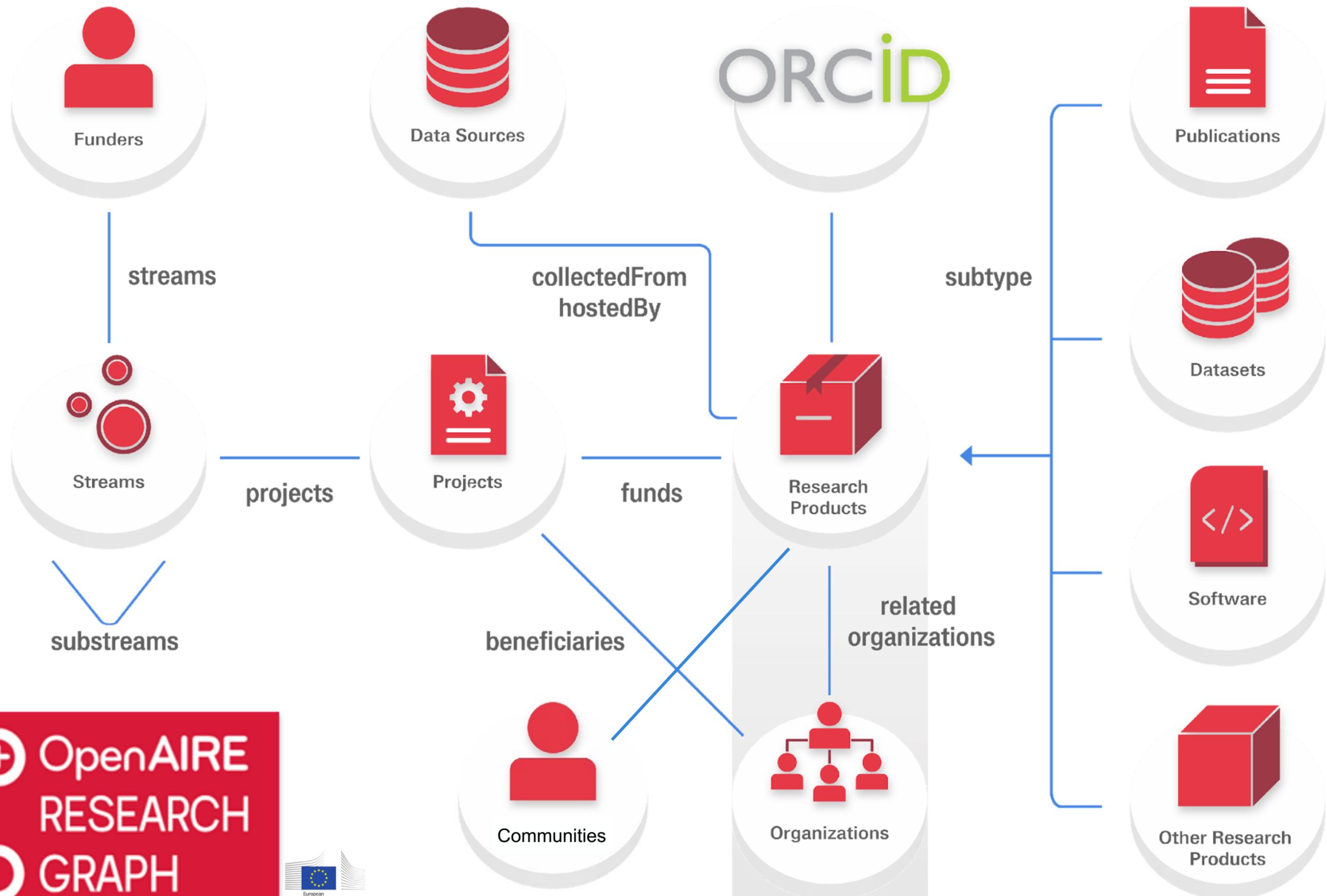


Discovery Open Science

Discovery of reproducible science outcomes must find new ways, driven by "scientific intentions" that go beyond the "find articles related to a research topic".

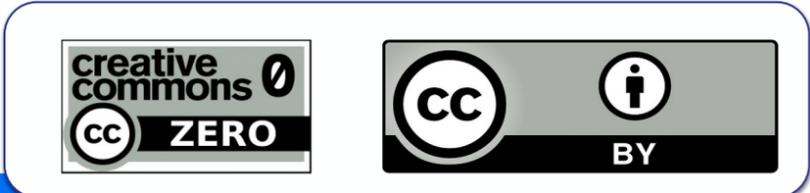


OpenAIRE Resaerch
Graph
*Tracking Open
Science*

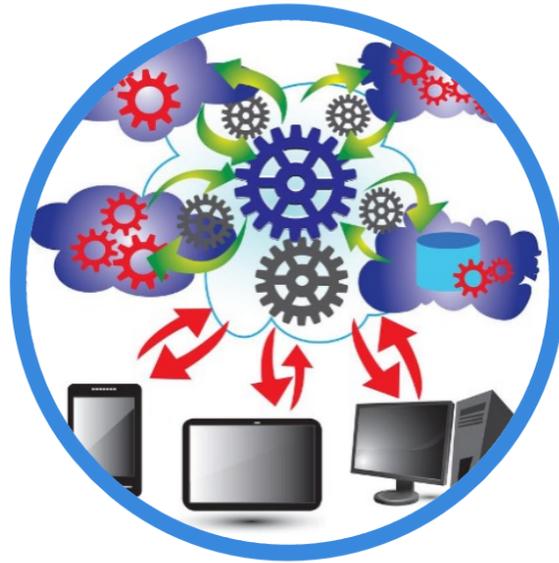




Providing an **open metadata research graph of interlinked scientific products, with Open Access information, linked to funding information and research communities**



Data wrangling challenges



HETEROGENEITY

Standards, crosswalks,
community and cross-
community view

[GUIDELINES.openaire.eu](https://guidelines.openaire.eu)



QUALITY & ACCURACY

Trust and stability of the data:
researchers and operators
in the loop

PROVIDE (validation,
brokering), user feedback,
PIDs



COMPLETENESS

Practices, sharing of technology
and data (Knowledge Graphs)

Open Science publishing
workflows, RDA OSG
interoperability

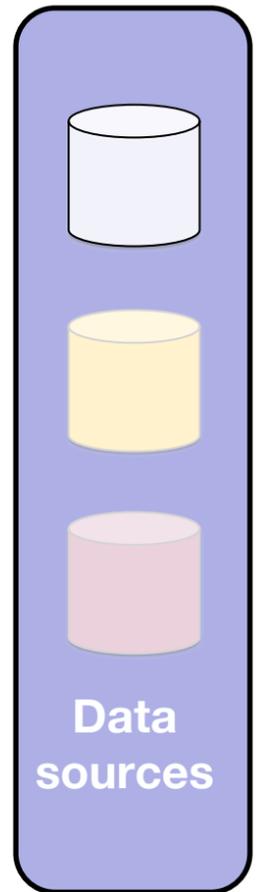
OpenAIRE Research Graph: the supply chain

Figures (Apr 2022-today):

- Direct data sources +2400
- Products +360Mi
- Harvested/mined links 480Mi +
- Full-texts 14.Mi+

Full-text mining 13 algorithms:
dataset, project, research-initiatives,
patent, software, pdb, affiliations,
communities, concepts, covid-19,
citations, similarity, classification

Propagation of properties via
relationships



Aggregation



Deduplication



Enrichment



indexing



The Dataverse Project

eprints repository software

DSpace

DataCite
FIND, ACCESS, AND REUSE DATA

euro CRIS

Different records representing the
same entity (results or organization)
are merged in one

Metadata records corresponding to
equivalent objects are merged.
Pre-print, post-print, published
versions are considered equivalent
for stats & monitoring purposes

Figures (Apr 2022):
500Mi/y API hits
650K/y user visits

Figures (Apr 2022):

Data sources 96K+ Funders 24
Publications 132Mi+ Projects 2.8Mi+
Datasets 16.5Mi+ Organizations 169K+
Software 270K+
Others 5.3Mi+



EC Participant Portal



Features



Provenance

Data sources
(record level)

Inference by Mining
(property level)



Metrics

Record level
Dedup level



Deduplication

Merge and
distribution of
relationships



PID agnostic

DOI, ArXiv, PDB,
PubMedID,
handle, etc.

PIDs in OpenAIRE

Authors

ORCID



Organizations

GRID

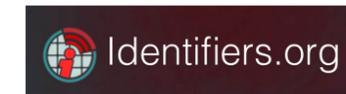
ROR



Products



arXiv.org



Funders/Projects



FCT Fundação para a Ciência e a Tecnologia



and other funders



Data sources

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

DOAJ
DIRECTORY OF OPEN ACCESS JOURNALS

OpenDOAR

+ journal lists from publishers



Services (fortcoming...)



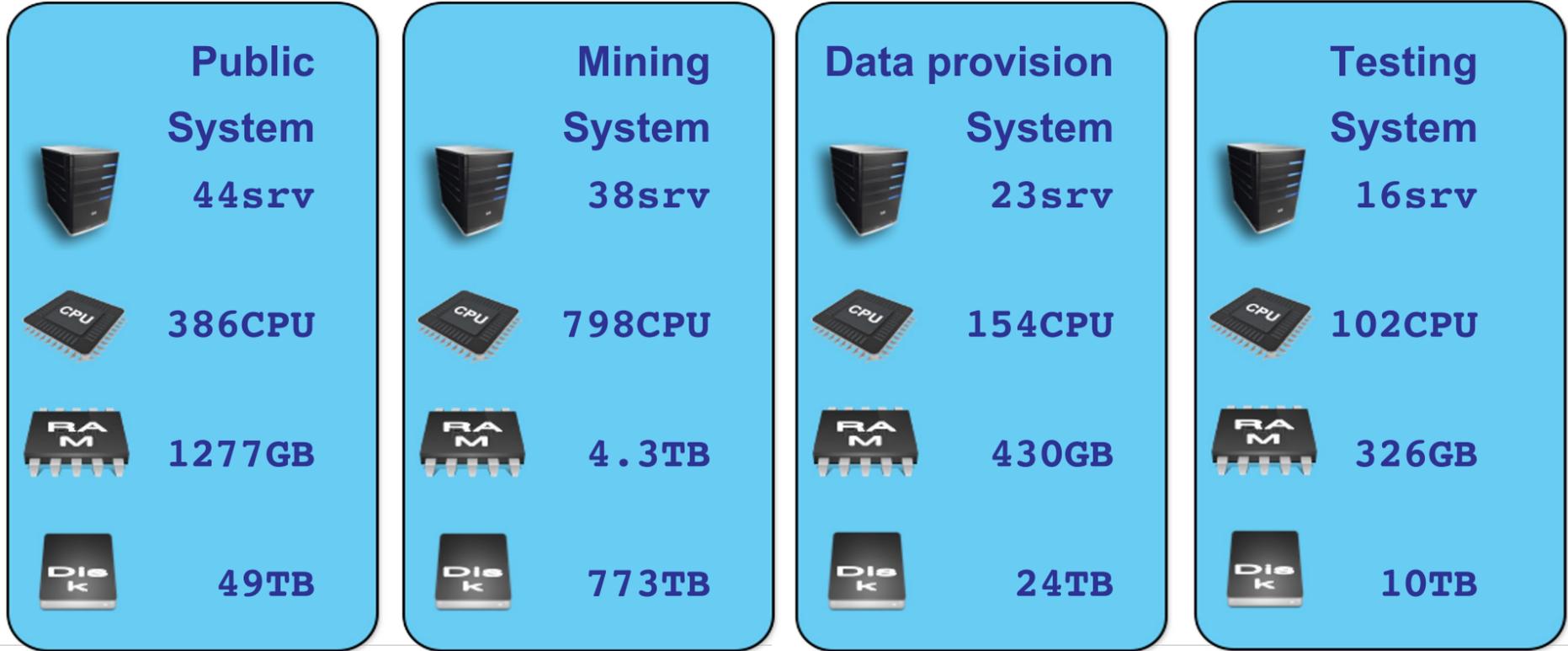
...Clusters and RIs service catalogues

OpenAIRE Research Graph: PDFs for full-text mining

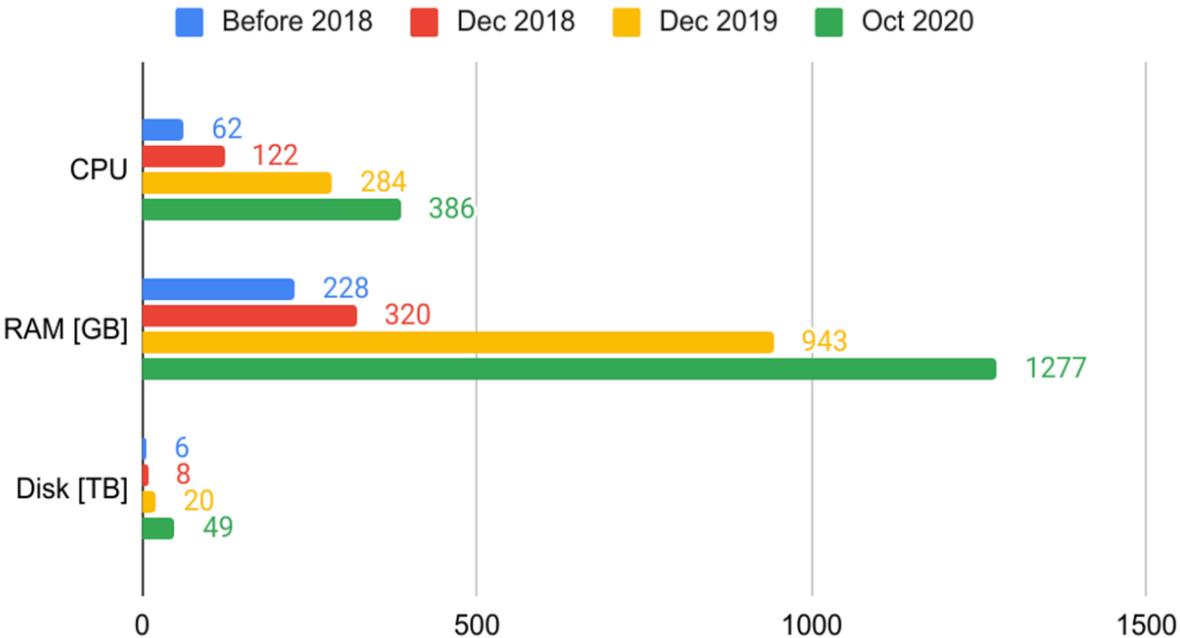
Publication full-texts
14.Mi+



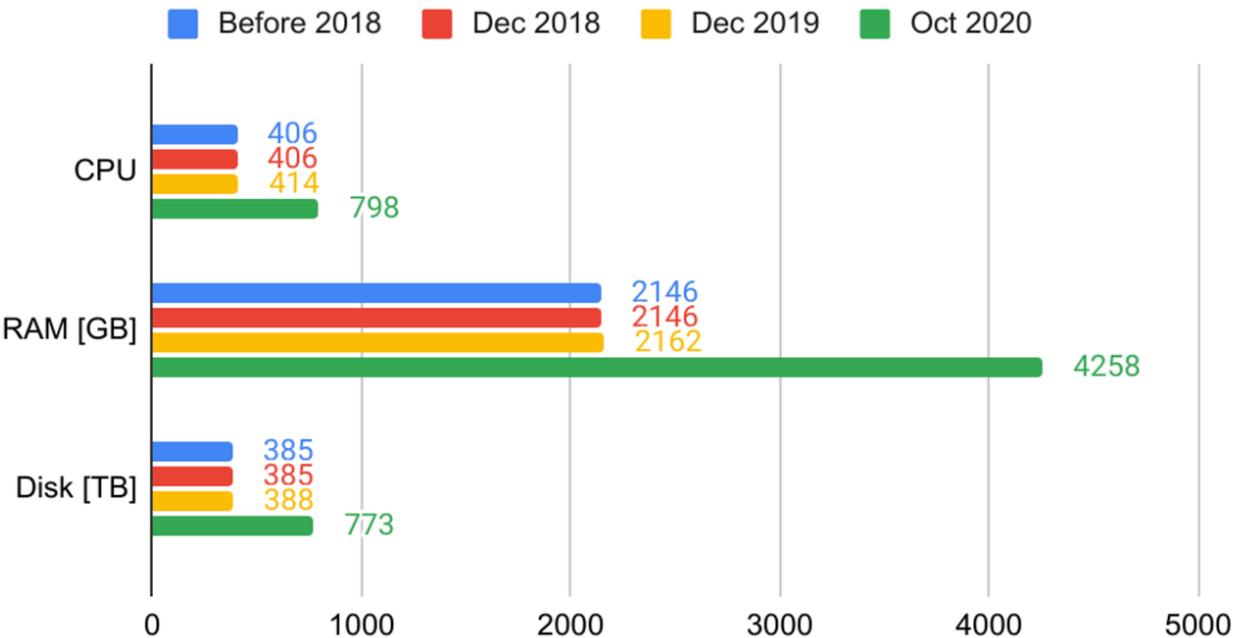
System infrastructure



Public subsystem resources



Mining subsystem resources



APIs and access

Access points

REST APIs with tokens
Linked Open Data
Dumps in Zenodo

Numbers

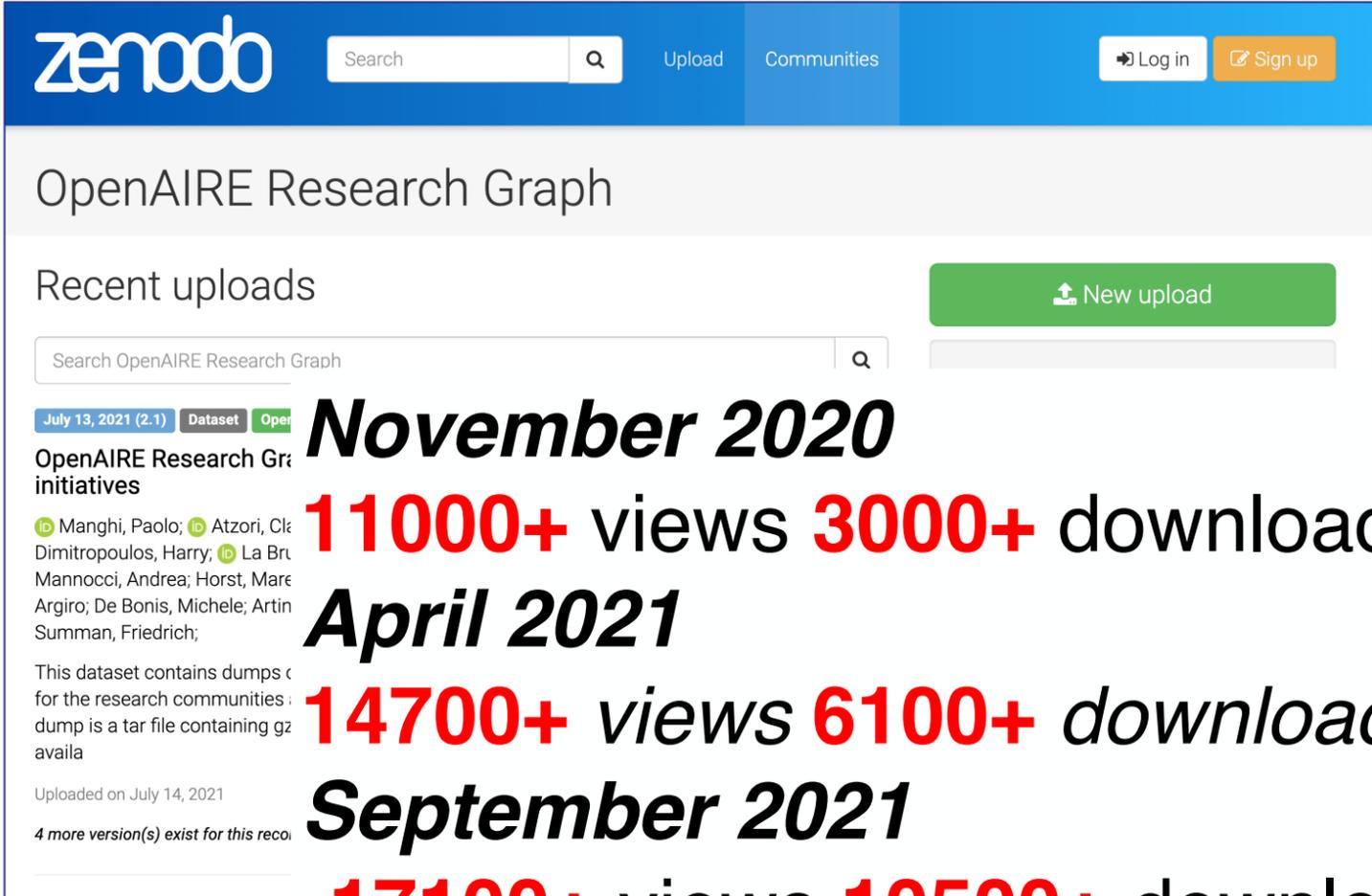
Average API hits
500Mi per year

<http://develop.openaire.eu>

OpenAIRE Graph Dumps

DEVELOP

- OpenAIRE Research Graph
- OpenAIRE COVID-19
- OpenAIRE Communities
- OpenAIRE LOD
- DOIBoost
- Scholexplorer 
- OpenAIRE RDF PID Graph



zenodo Search Upload Communities Log in Sign up

OpenAIRE Research Graph

Recent uploads [New upload](#)

Search OpenAIRE Research Graph

July 13, 2021 (2.1) Dataset Open

OpenAIRE Research Graph initiatives

Manghi, Paolo; Atzori, Cl...
Dimitropoulos, Harry; La Bru...
Mannocci, Andrea; Horst, Mare...
Argiro, De Bonis, Michele; Artin...
Summan, Friedrich;

This dataset contains dumps c...
for the research communities :...
dump is a tar file containing gz...
availa

Uploaded on July 14, 2021

4 more version(s) exist for this reco...

November 2020

11000+ views **3000+** downloads

April 2021

14700+ views **6100+** downloads

September 2021

17100+ views **10500+** downloads

April 2022

21240+ views **26305+** downloads

<https://zenodo.org/communities/openaire-research-graph>

Consumers

DEVELOP



The Graph as EOSC Resource Catalogue



ORCID

Early 2021

Connecting Research and Researchers



CORDIS

Grant Management | Project Continuation Report

Publications

Suggested publications from OpenAIRE (2 publications)

No.	Type	Title	Authors	Title of the Journal/Proc./Book	Date of acceptance	DOI	Request	actions
1	Book/In	Open Science Training Handbook	Senja Berjaks, April Elyt		04/04/2018	10.5201/zenodo.121149 10.5201/zenodo.121147		X
2	Other	Project FOOTER Plus	Drinken, Helene		07/09/2017	10.5201/zenodo.88720 10.5201/zenodo.88722		X

Project publications (1 publications)

No.	Type	Title	Authors	Title of the Journal/Proc./Book	Number, date or issue of the Journal/Proc./Book	DOI	Request	actions
1	Publication	Project FOOTER Plus	Drinken, Helene			10.5201/zenodo.101328		X

Annotations: "Accept the suggested publications from OpenAIRE and add" (pointing to the suggested list), "your project publications" (pointing to the project list), and "Add" (pointing to the bottom right button).





Call for collaboration

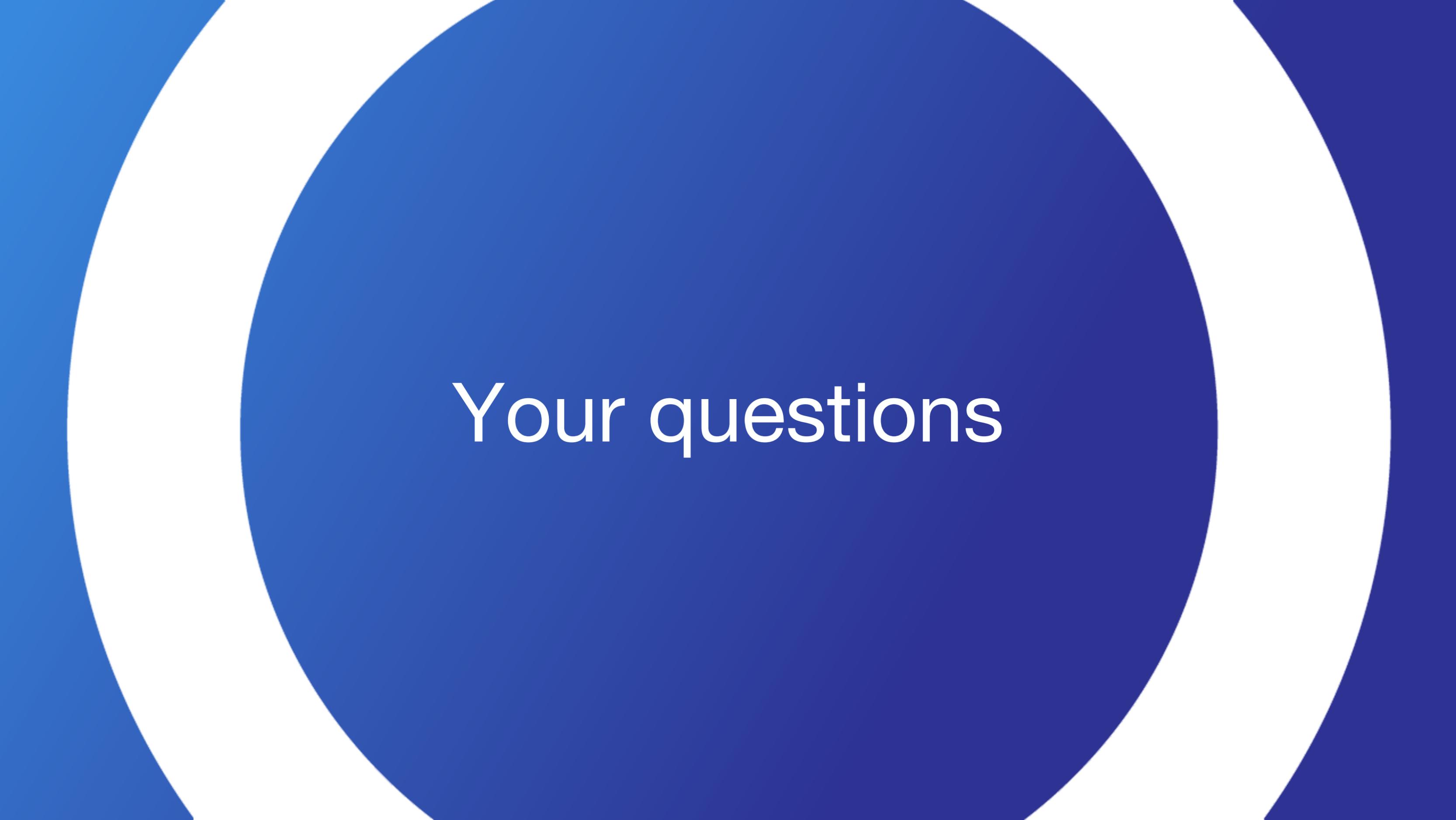
Improving quality and completeness the Graph

Data and infrastructure sharing
OpenAIRE support to access to Graph data (metadata and full-text) and access to full-text mining infrastructure

Embedding improvements as part fo the data provision workflow of the Graph

Devising new Open Science services for the community

Improving service provision



Your questions

Thank you!

Paolo Manghi (OpenAIRE/CNR-ISTI)

paolo.manghi@openaire.eu

graph.openaire.eu

