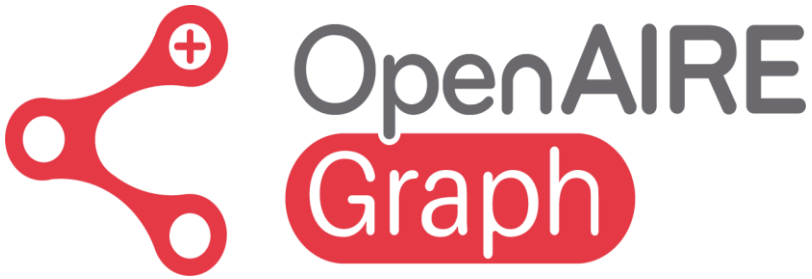
@openaire_eu

The OpenAIRE Graph: *supporting discovery and monitoring of Open Science research with open data*

Paolo Manghi
OpenAIRE AMKE

The OpenAIRE Graph: the basics

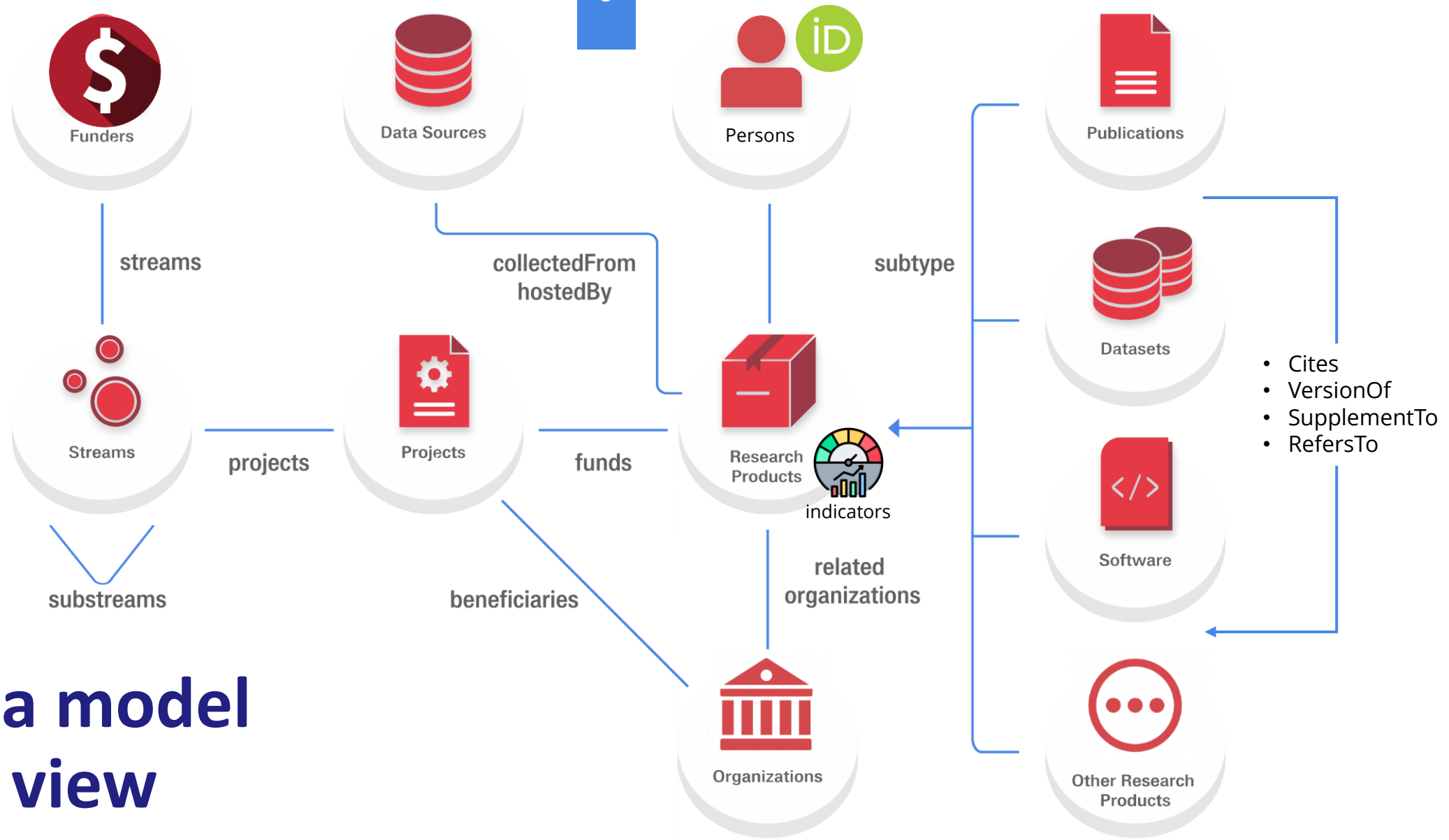


Scholarly Knowledge Graph

A collection of metadata describing entities of the research lifecycle and relationships among them

- Open
- Complete
- De-duplicated
- Transparent
- Participatory
- Decentralized
- Trusted

The data model moon's view



Some data model features

Embeds indicators such as citations (from OpenCitations Meta, MAG, DataCite, and OpenAIRE), APCs (from OpenAPC), COUNTER metrics (from OpenAIRE UsageCounts), and other indicators (e.g. popularity and influence from BIP!)

Persistent identifiers include today DOI, ArXiv, PubMed, Handles for *publications;* DOI, accession numbers, handles for *research data;* DOI, handles, and software heritage IDs for *research software* (the data model can flexibly include any PID schema!)

Stable identifiers for graph records

Provision chain

Onboarded data sources

OpenAIRE Interoperability Guidelines

Repositories | Publishers
OA Journals | Registries
Aggregators | CRIS

Notable examples:

zenodo | EPIsciences

PROVIDE

Metadata Validator

OpenAIRE BROKER

Instrumental data sources

From OpenAIRE partners:

OPEN ORGS | OpenCitations | open APC

Funder databases

SciNoBo | BIP! DB | OpenAIRE UsageCounts | DOAJ

Microsoft Academic | ORCID | unpaywall

DataCite | EMBL-EBI | PubMed | Crossref

full - texts

Aggregation → Enrichment by mining → Deduplication → Enrichment by inference → Finalisation

publication – project/funder links
publication – data/software links
publication – publication citations
publication – organization affiliations
product – product similarity
FOS, SDGs
....

Fields of Science

SUSTAINABLE DEVELOPMENT GOALS

eosc EU Node

OpenAIRE Graph

Public Graph Dataset & API

Data sources contributing to the Graph

Spotlights on data acquisition
2.5k+ direct data sources (140K+ total)

GitHub
Applications Database
Software Heritage
elixir bio.tools
… and more **tools and software sources**

frontiers
PLOS
Springer
F1000 Faculty of 1000
… and more **publishers**

CORDIS
NSF
DFG Deutsche Forschungsgemeinschaft
… and more **European and international funders**

unpaywall
Microsoft
OpenCitations
… and more **research s**

OpenAIRE Graph

DataCite
re3data.org REGISTRY OF RESEARCH DATA REPOSITORIES
ROR
Crossref
OpenDOAR
FAIRsharing.org standards, databases, policies
ORCID
… and more **registries**

LA Referencia Red de repositorios de acceso abierto a la ciencia
SciELO
DOAJ DIRECTORY OF OPEN ACCESS JOURNALS
… and more **aggregators**

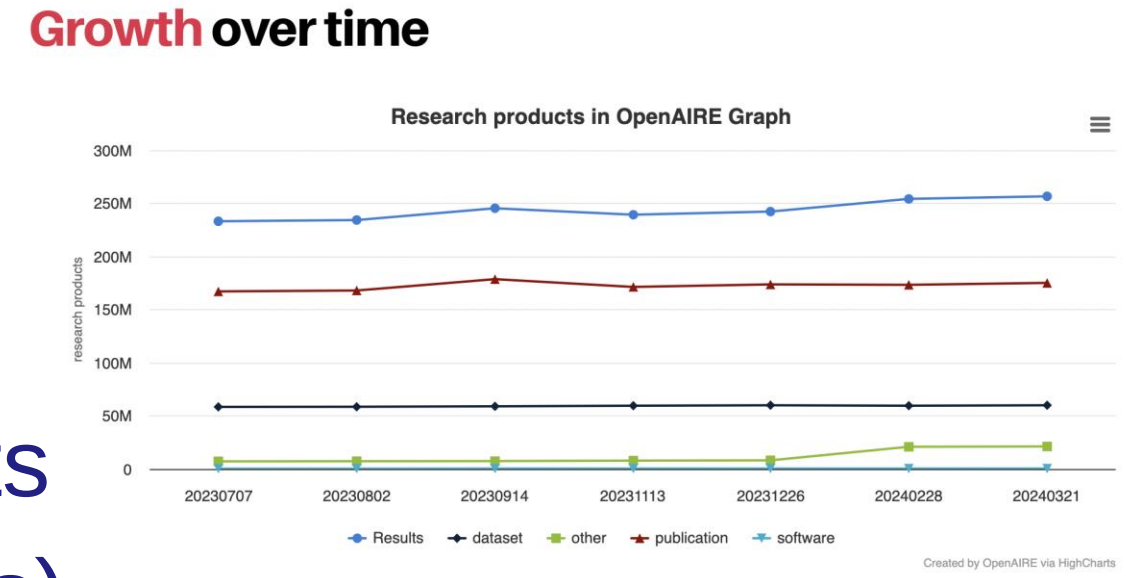
Cornell University Library
arXiv.org
RePEc
PubMed
HAL archives-ouvertes.fr
zenodo Research. Shared.
… and more **thematic and institutional repositories**

DARIAH-EU
elixir
EPOS EUROPEAN PLATE OBSERVING SYSTEM
EOSC-Life
ECRIN EUROPEAN CLINICAL RESEARCH INFRASTRUCTURE NETWORK
egi
… and more **einfra and RI sources**

OpenAIRE

OpenAIRE Graph in numbers (25th of Oct., 2024)

- 194Mi publications
- 405k software
- 64Mi research data
- 23Mi other research products
- 186 funders (30 with projects)
- 3.4Mi grants
- 318K organizations

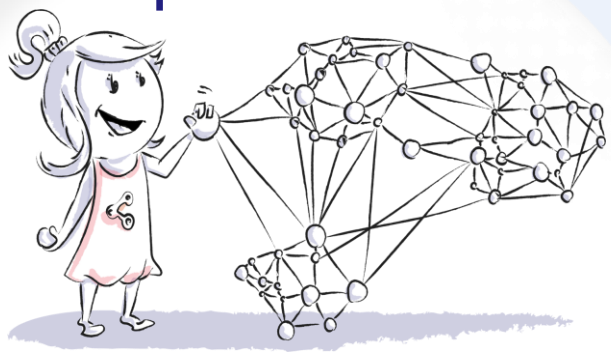


OpenAIRE Graph

OpenAIRE

OpenAIRE Graph in numbers

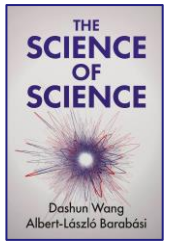
- **Links from pubs to products**
  - 2.1Bi pubs-pubs links (citations)
  - 138Mi pubs-data links
  - 437K pubs-software links
- **Links to grants**
  - 5.4Mi pubs-grant links
  - 862K data-grant links
  - 16.3K software-grant links
- **Affiliations**
  - 172Mi pubs-organization links (affiliations)
  - 3.7Mi data-organization links (affiliations)
  - 166K software-organization links (affiliations)
- 432Mi pubs-FoS (77Mi pubs with FoS)

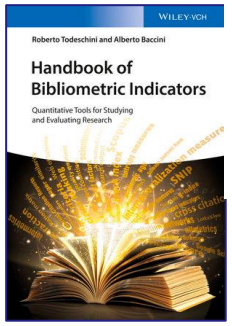


OpenAIRE

Consumers

Researchers





The MIT Press Journals

QSS

"Transparent peer review will serve as a valuable tool for readers to see how quality control is done at *Quantitative Science Studies.*"

– Ludo Waltman, Editor-in-Chief

Service Providers/Companies

eosc

Scopus

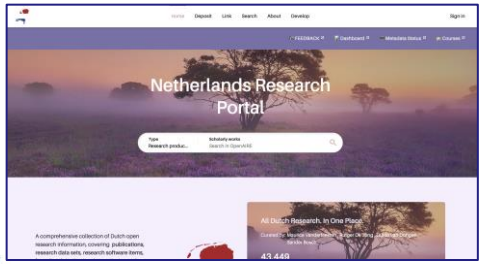
SPRINGER NATURE

DSPACE

Reporting to funders,
Discovery of pubs, data, software
Assessment, monitoring (FAIRness,
Open Science, Open Access)

Organizations

NATIONAL OPEN ACCESS MONITOR



THE WORLD BANK

OpenAIRE

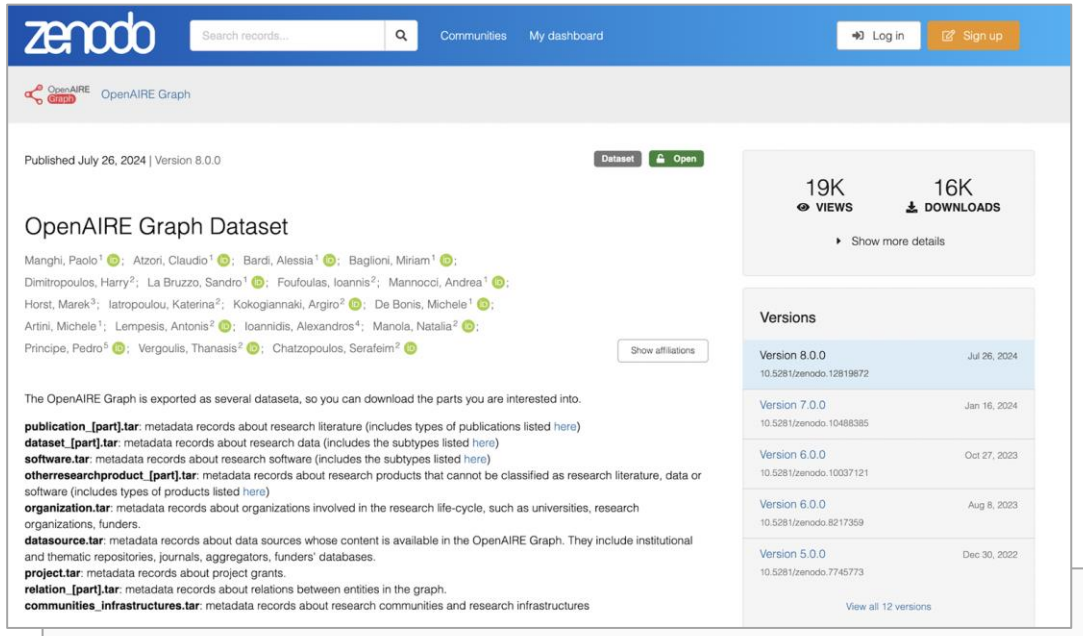
Funders

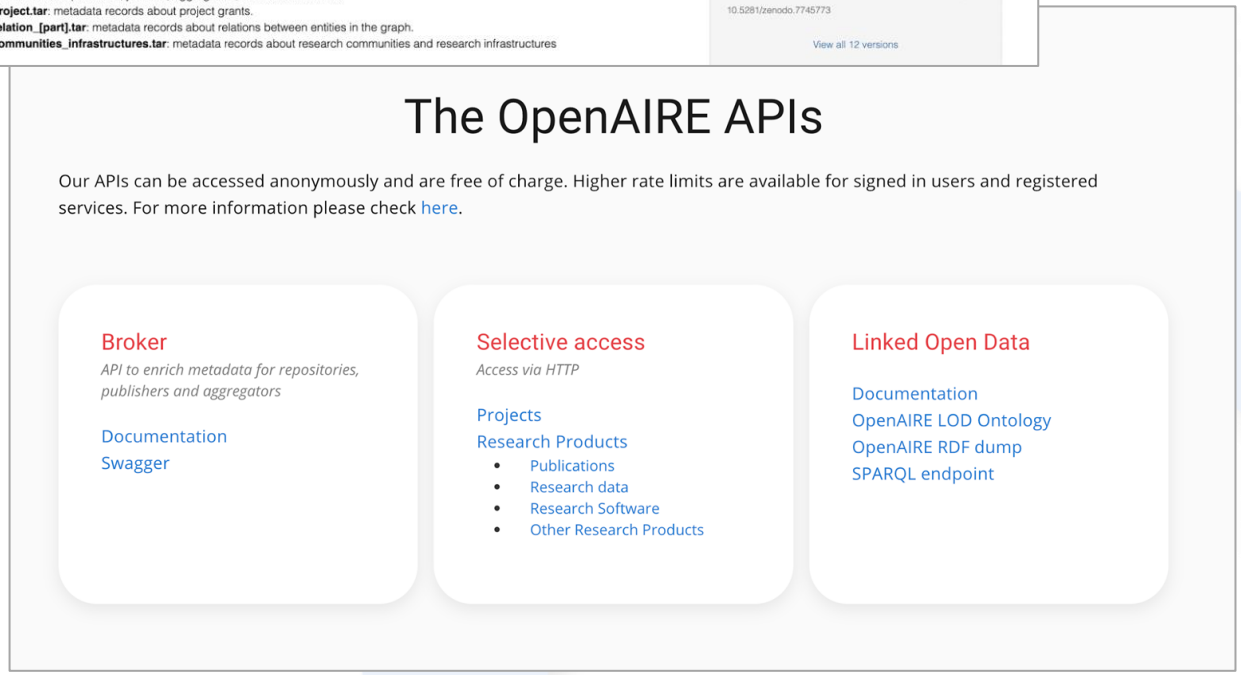
European Commission

FCT
Fundação
para a Ciência
e a Tecnologia

Access to the Graph

- Dataset in Zenodo
- API on documentation web site (graph.openaire.eu/docs)



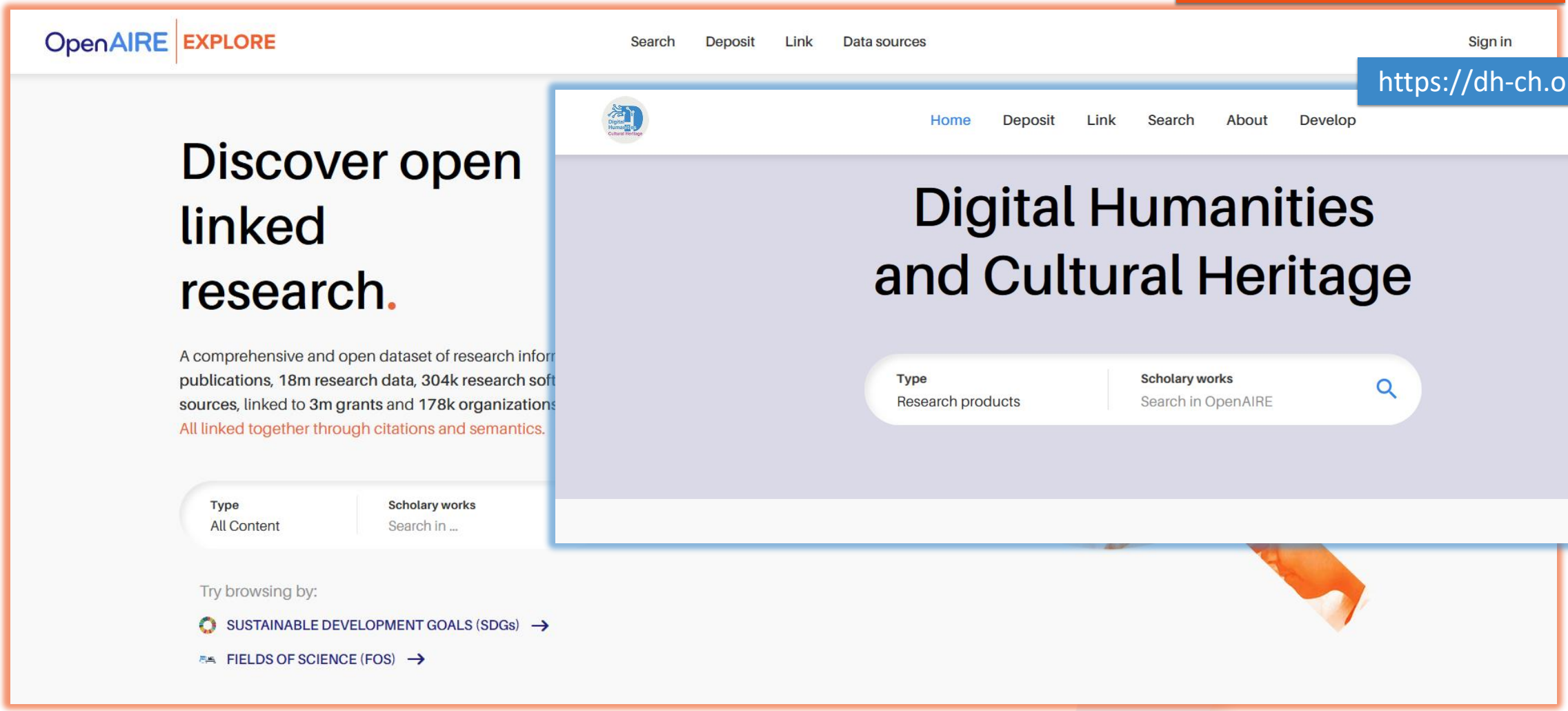


OpenAIRE

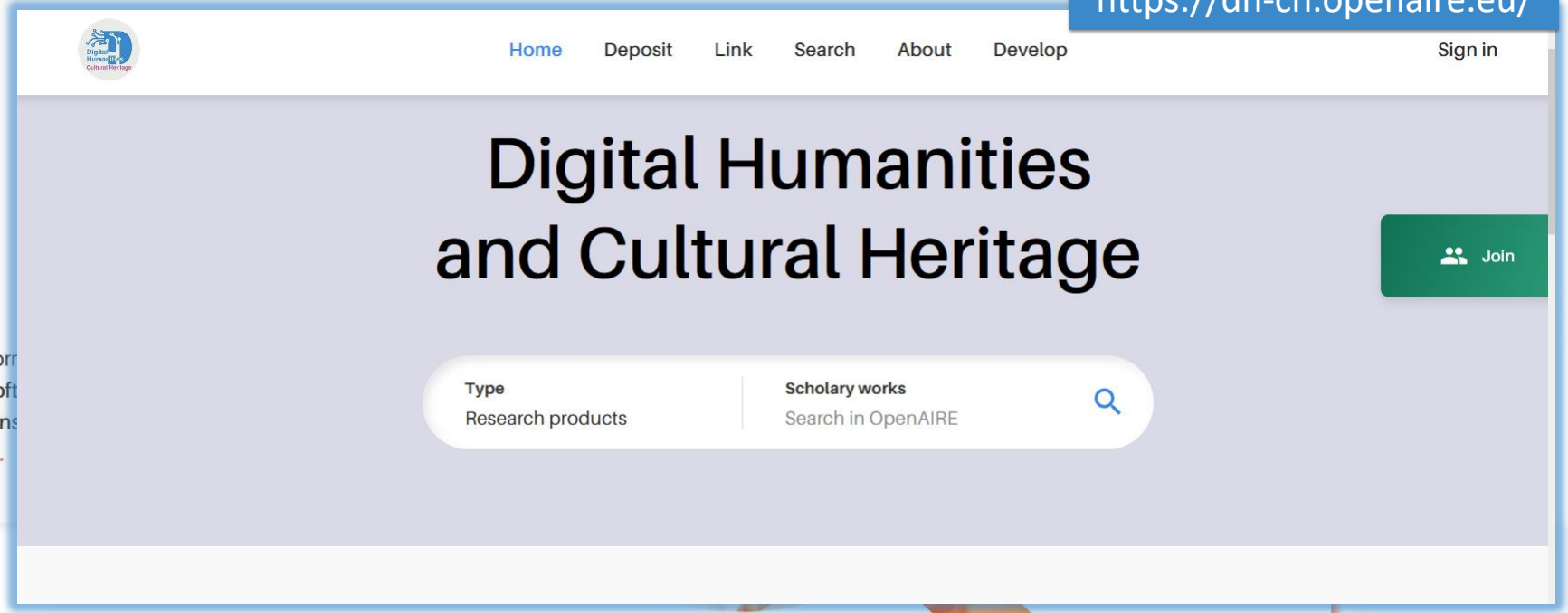
Applications: discovery and research assessment

OpenAIRE EXPLORE and CONNECT

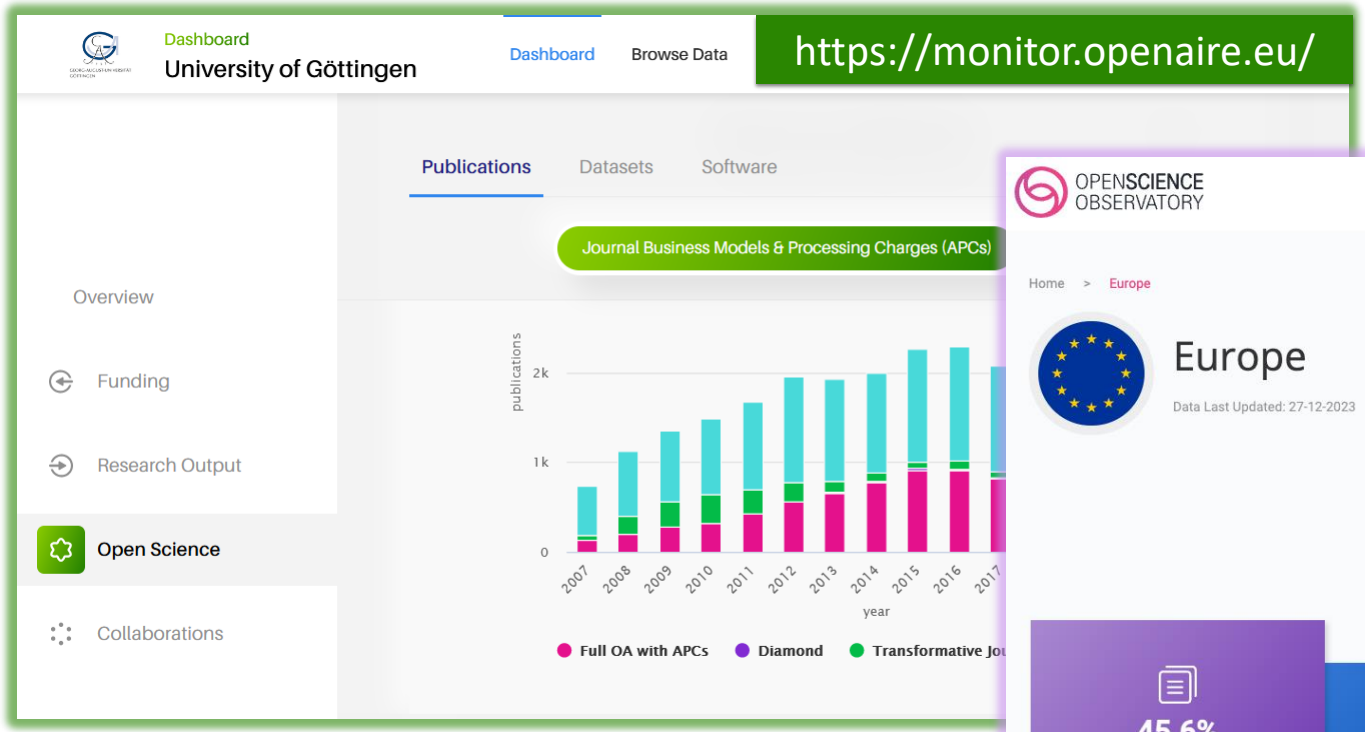
https://explore.openaire.eu/

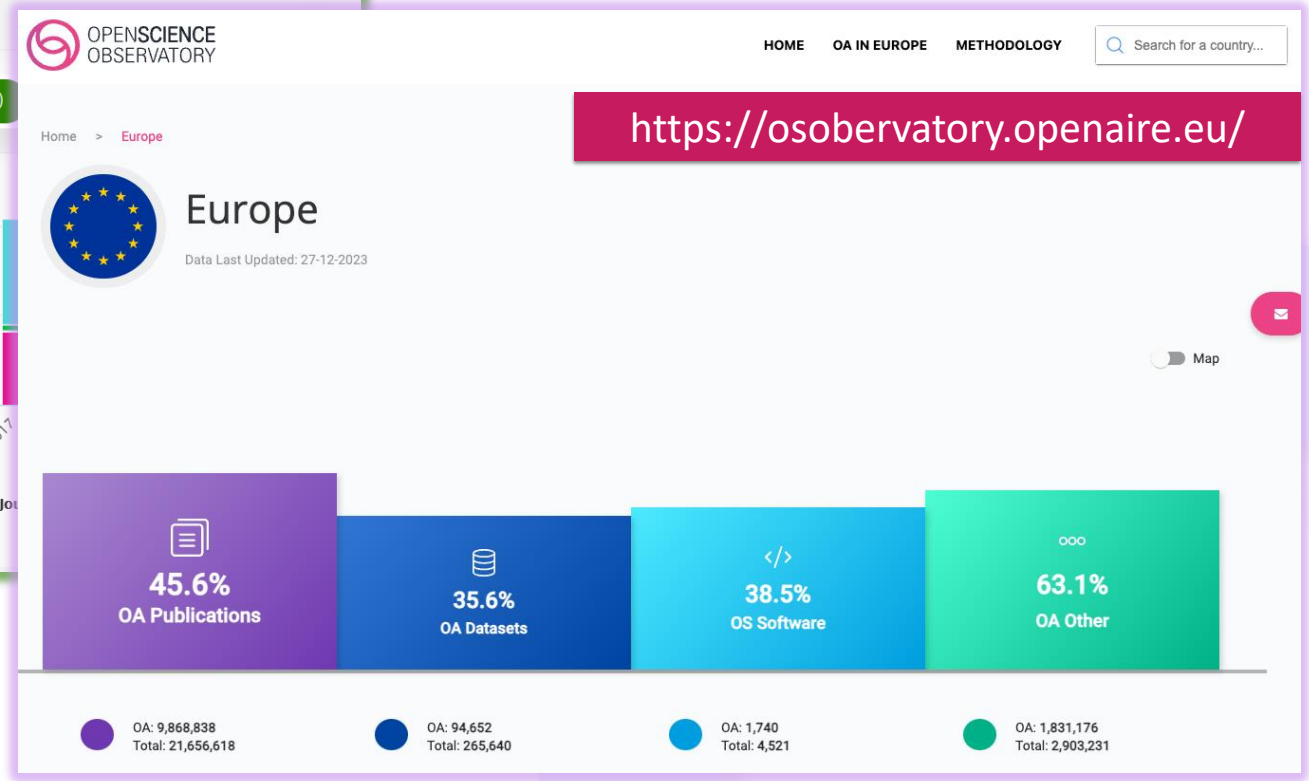


https://dh-ch.openaire.eu/



OpenAIRE MONITOR and OSmonitor

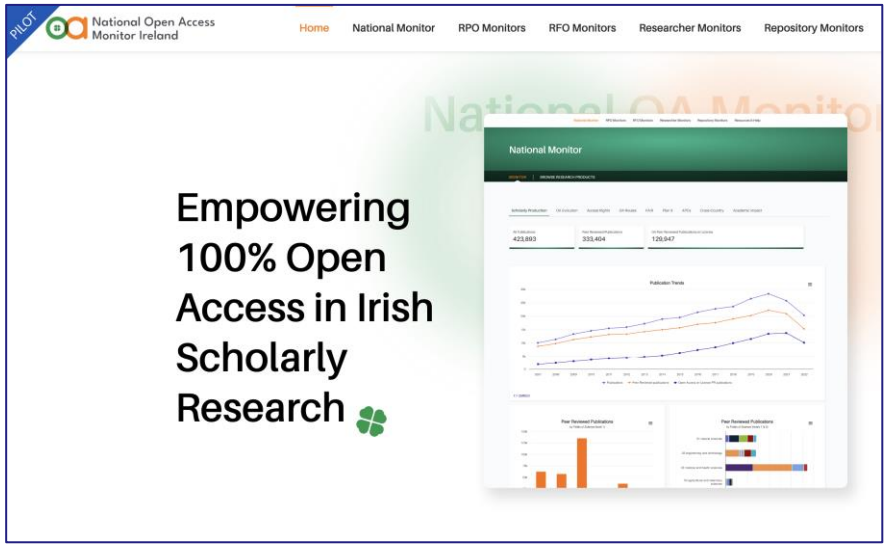




OpenAIRE

The Irish OS monitor tender

https://oamonitor.ireland.openaire.eu



Onboarded data sources

OpenAIRE Interoperability Guidelines

Repositories | Publishers
OA Journals | Registries
Aggregators | CRIS

Notable examples:

zenodo | EPIsciences

PROVIDE

Metadata Validator

OpenAIRE BROKER

Instrumental data sources

From OpenAIRE partners:

OPEN ORGS | OpenCitations | open APC

Funder databases

SciNoBo | BIPI DB | OpenAIRE UsageCounts | DOAJ

Microsoft Academic | ORCID | unpaywall

DataCite | EMBL-EBI | PubMed | Crossref

full - texts

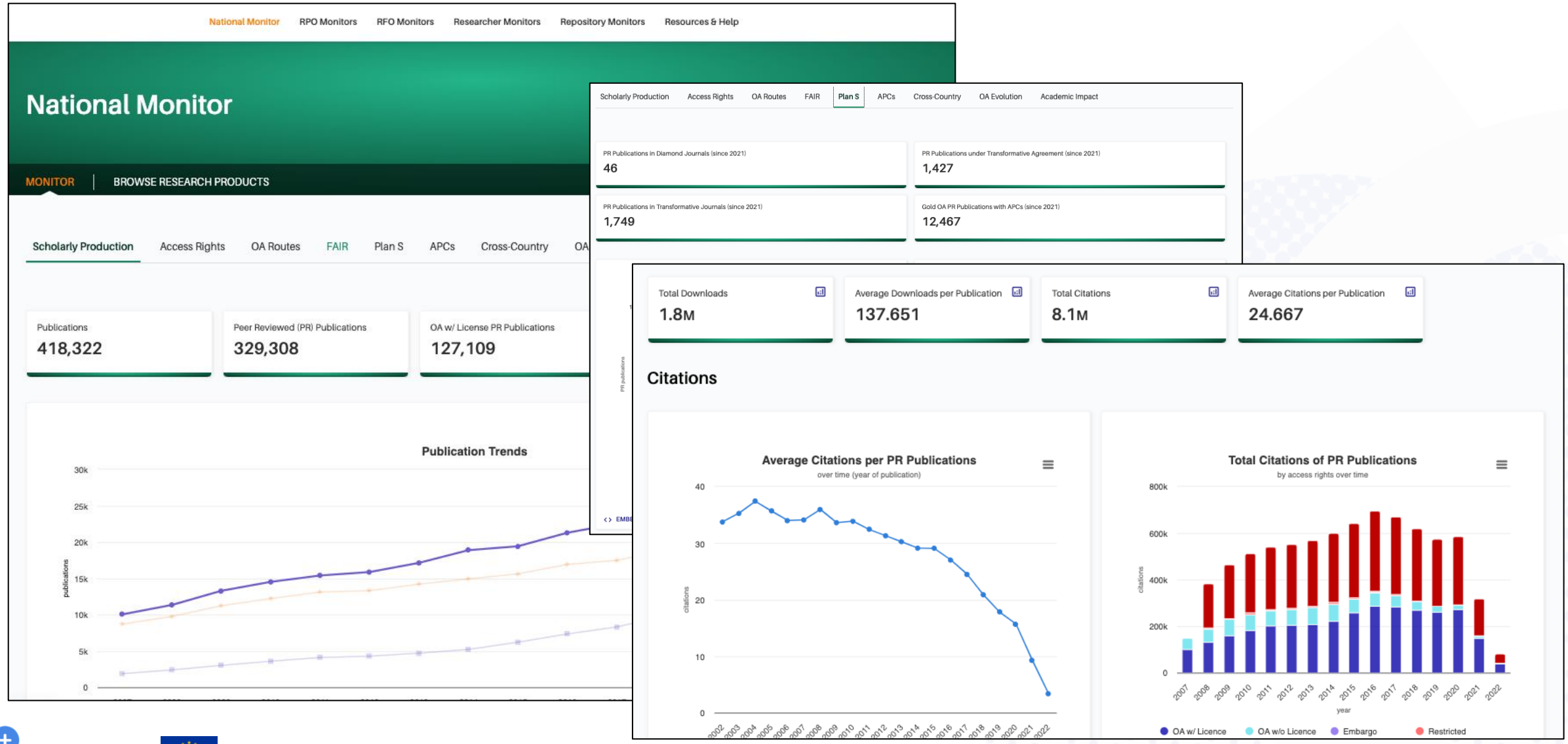
Aggregation → Enrichment by mining → Deduplication → Enrichment by inference → Finalisation

OpenAIRE Graph

Public Graph Dataset & API

User feedback

The Irish OS monitor tender



For and by the community

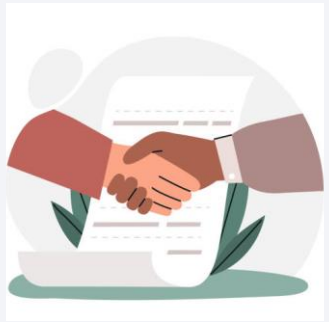
The Graph is public good!

In the hands of the Community

Community governance

Members of OpenAIRE can evaluate results and steer Graph's roadmap and vision by contributing to OpenAIRE Steering Committees

Members can join the technical team and contribute to the development and operation of the Graph



Community alignment

Synergies with repository managers, libraries, service provides at the global level to ensure alignment of data models, metadata, vocabularies at the global level (***EOSC guidelines, OpenAIRE Guidelines***, SKG interoperabilty framework)

Community involvement

Seeking collaborations to enrich quality and coverage of data

We're interested to integrate solid research results in the domain (e.g., SDGs inference)

We're interested to include high-quality data sources (e.g., OpenCitations integration)

End-users feedback

We care adapting roadmap to match community expectations, e.g. research organizations, funders, companies

We care about identifying and fixing errors

We acquire Graph data enrichment from users

Improving scholarly communication infrastructure and workflows

We provide tools for enriching data sources with metadata information they are missing

We provide tools to facilitate researchers, publishers, service providers in "publishing" research trusted scientific outcomes with high-quality metadata

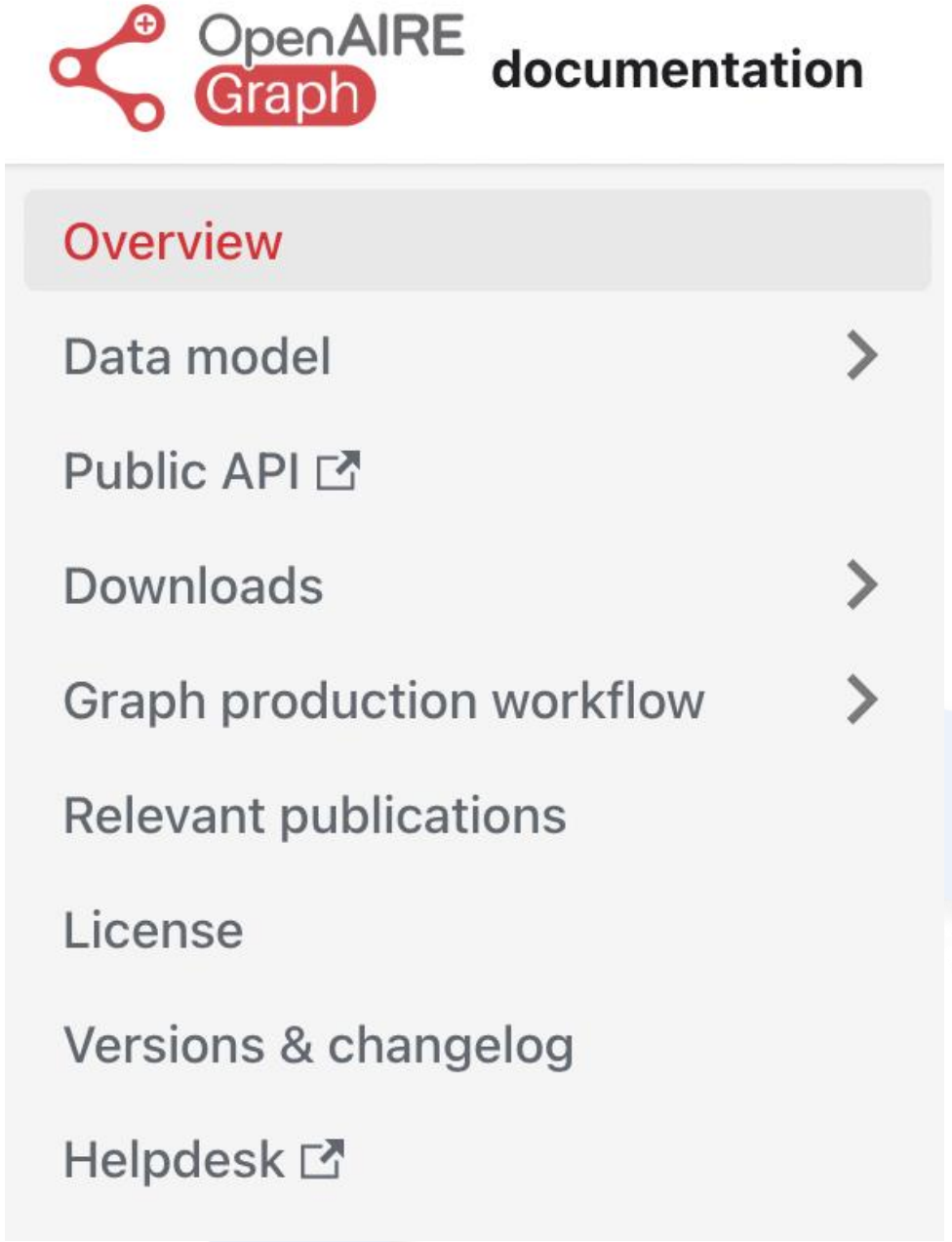
OpenAIRE Graph: take home message

It’s the largest Open Science-focused graph	Broader **coverage** of publications, research data, and research software, together with links between them
It’s open	Accessible via APIs and datasets in Zenodo under **CC-0**
It’s transparent	We document the process of generation and track **provenance**
It’s a public good	A **public infrastructure** (like highways)
It’s participatory	**Welcomes data/methods contributions** from service providers, organizations, and researchers

For transparency

`https://graph.openaire.eu/docs/`

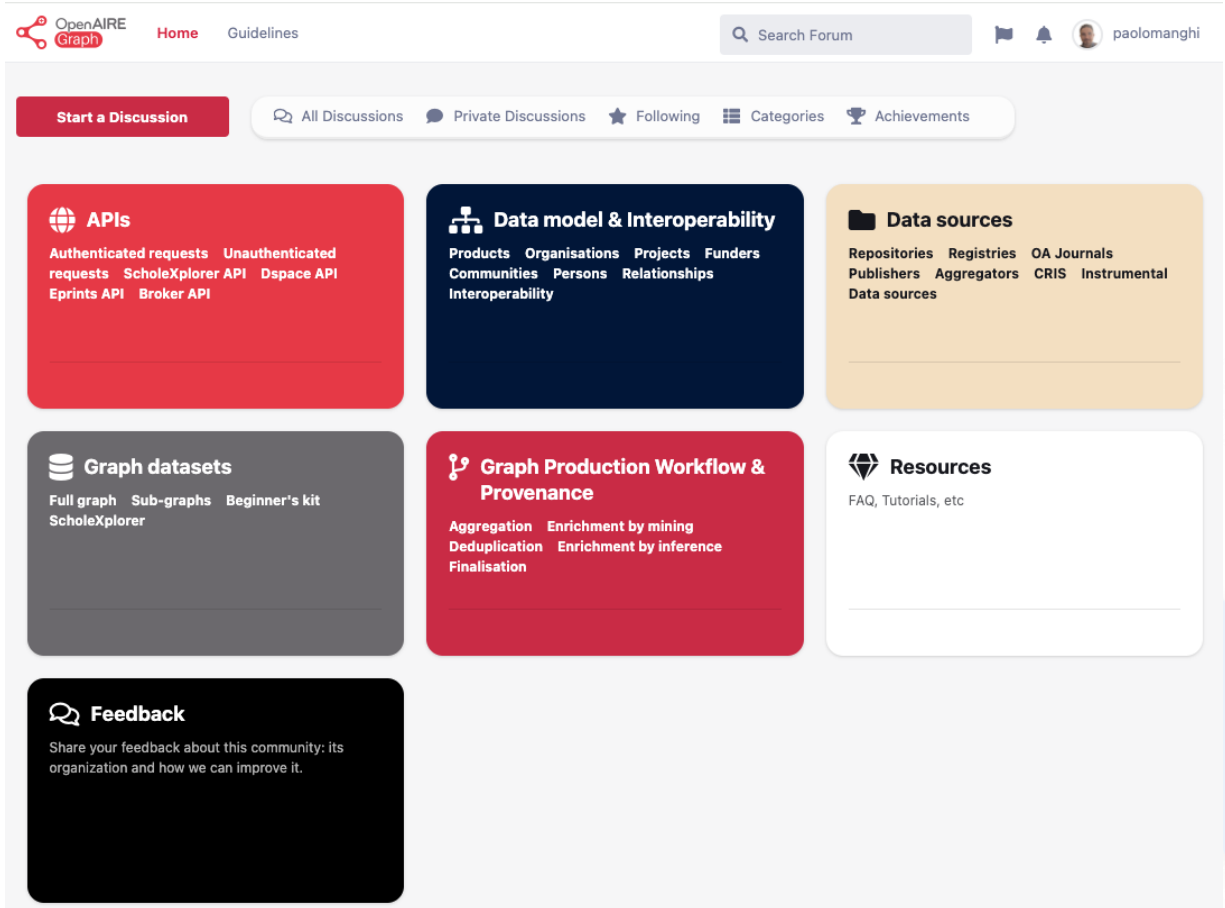
Documentation



For enquiries

`https://openaire.flarum.cloud`

Forum



OpenAIRE

Thank you for your attention!

paolo.manghi@openaire.eu