



OpenAIRE Graph gettare le basi per l'analisi strategica della ricerca aperta



Paolo Manghi

Istituto di Scienza e Tecnologie dell'Informazione - CNR
OpenAIRE AMKE - Infrastruttura per la Ricerca Europea

OpenAIRE Graph

A Knowledge Graph of Research Information



Metadata knowledge graph

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

Timely and comprehensive coverage of research outputs

Monthly updates



Precision, depth & processing

Rigorous cleaning, deduplication, and enrichment for optimal accuracy

Full-text mining of links: Research results to projects, author affiliations, and classifications

AI methods for the identification of FoS, SDG



Robustness & openness

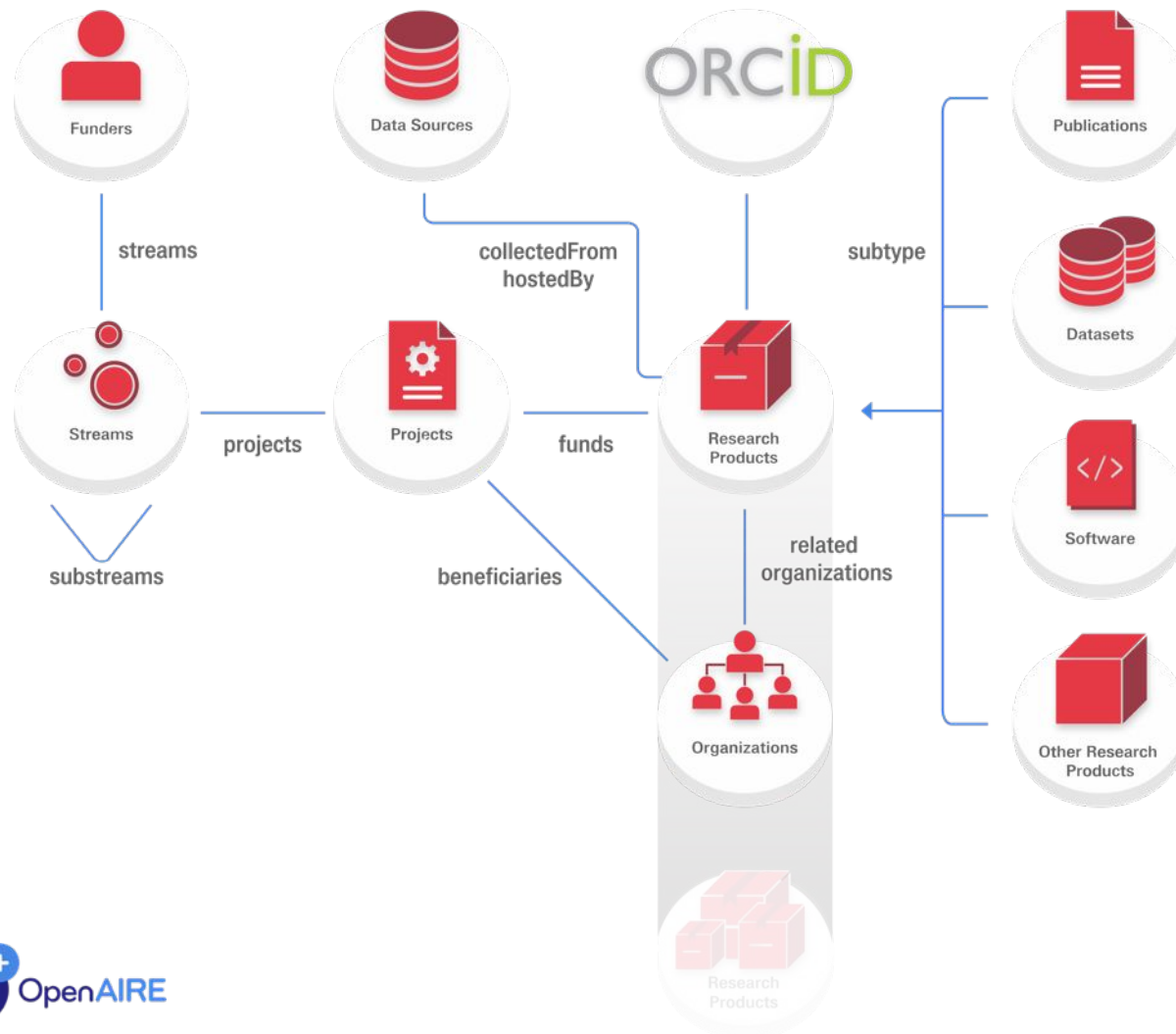
Professional infrastructure: maintenance, QoS, backups, overseen by the OpenAIRE technology centre (ICM)

Open data, Open Source & transparent methodologies

Beyond articles and publishers

“Scientific innovation is no longer only about peer reviewed literature”

OpenAIRE **Graph**: data model and main features



COAR Confederation of Open Access Repositories

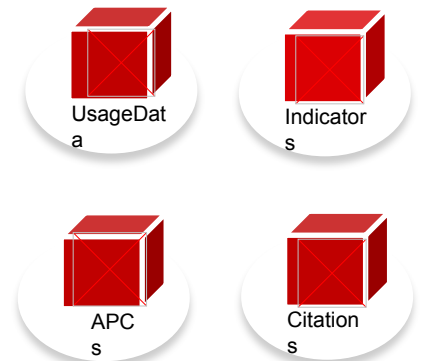
Beyond publications to include data, software, clinical trials, patents, etc.

Global & cross-disciplinary coverage

Enriched by **AI/FTM-processes** - citations, FOS, SDGs, domain specific terms/ontologies

Indicators-agnostic

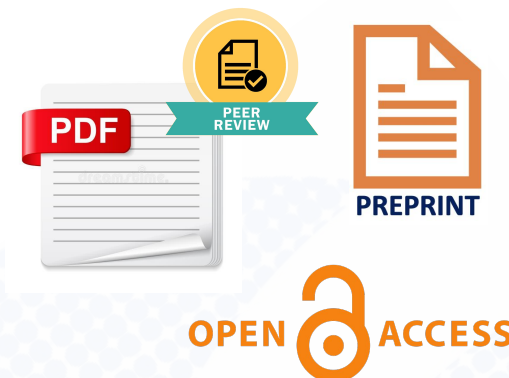
PID-agnostic



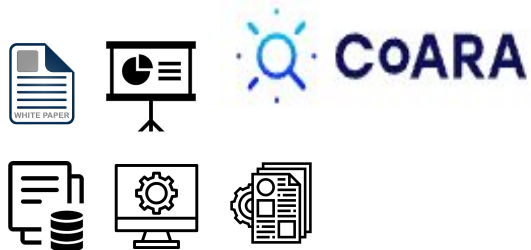
New classes of eligible research product typologies

Traditional typologies

Peer-reviewed publications, preprints, open access tags & pub-pub citations



Effective bibliotyping



Open Science typologies

Traditional typologies and dissemination products, research data, research software, patents & data/software citations

New classes of eligible publishing venues

“Adhering to community defined guidelines”

- Research data repositories
- Institutional repositories
- Catch-all repositories
- National & Thematic aggregators
- Diamond journals and CRIS systems
- Local registries of authors/organisations

Jisc

OpenDOAR

6,000

FAIRsharing.org
standards, databases, policies

2,000

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

3,500

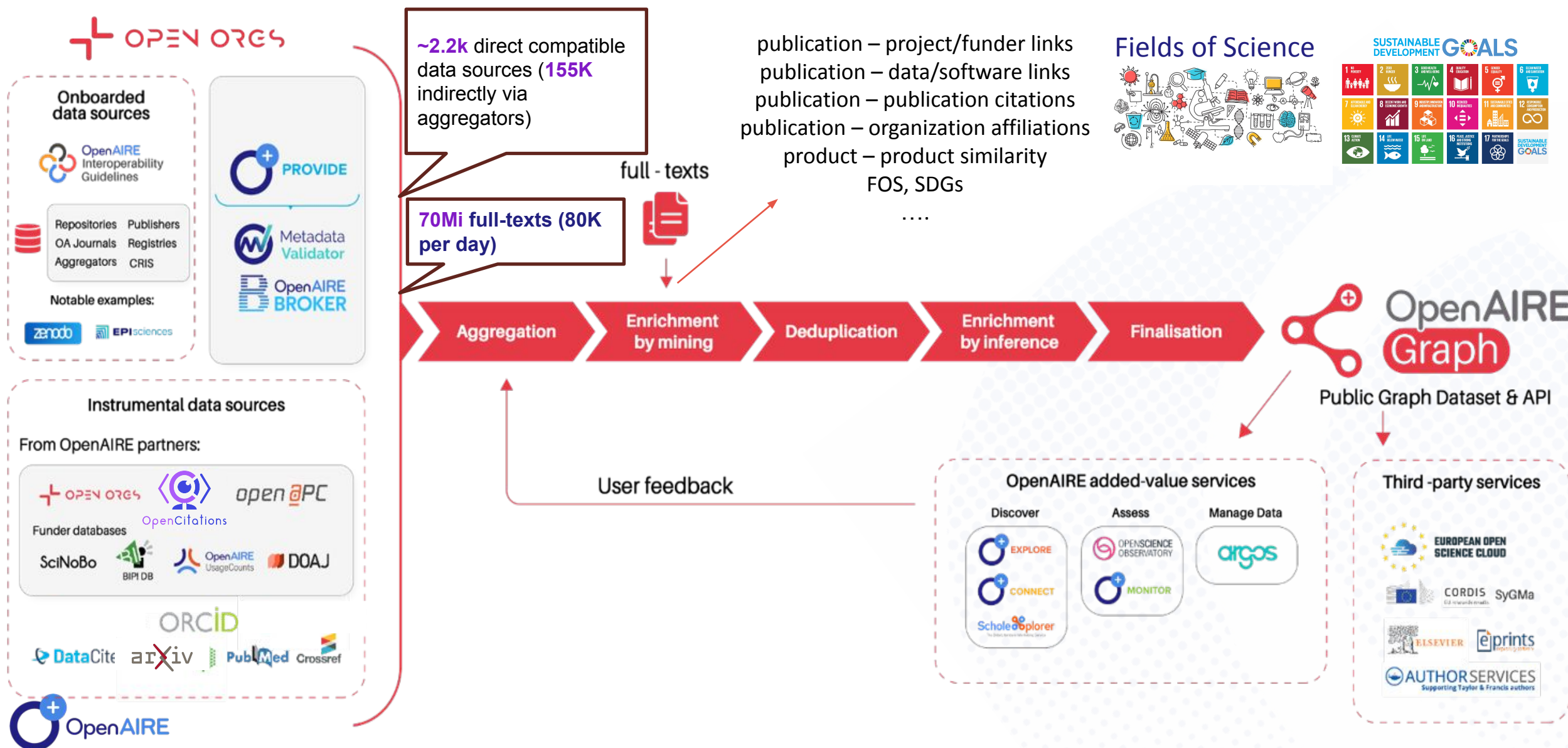
OpenAIRE **Graph** Data Sources



Metadata curation and quality

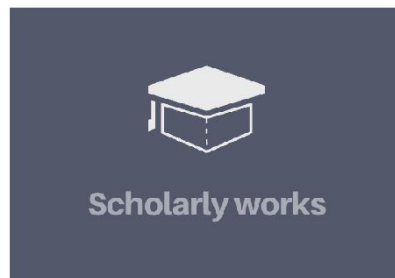
"Size matters, but without quality it's noise"

OpenAIRE Graph provision chain



OpenAIRE **Graph** entities

Latest Statistics



238,534,749

Publications



108,268,274

Data items



971,214

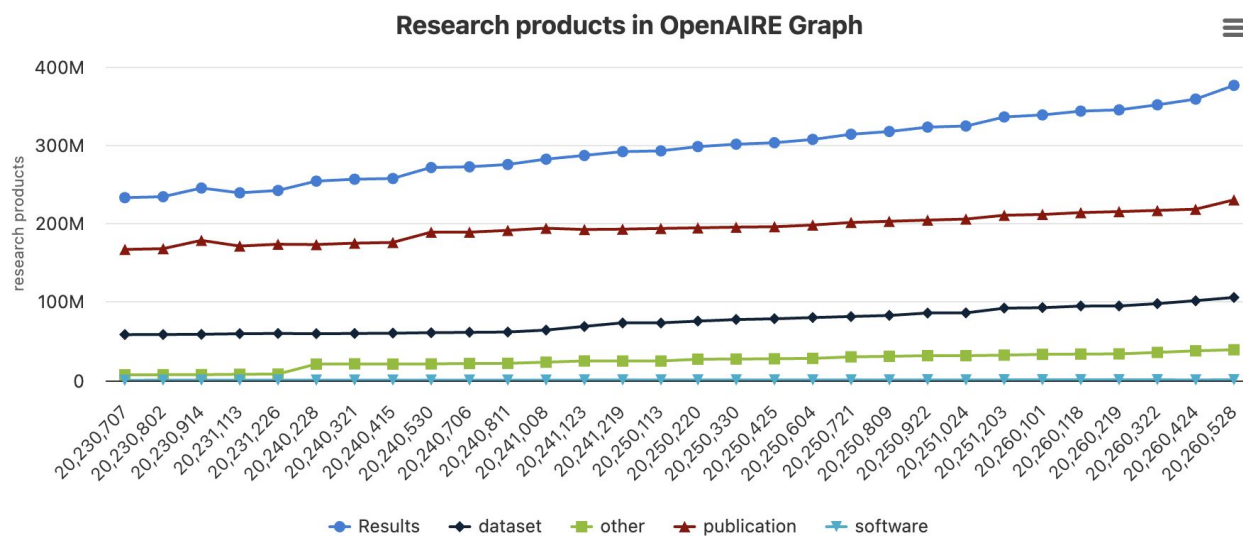
Software



42,208,140

Other

Growth over time



157,198

Data sources



324

Funders



3,924,351

Grants

390 M Research Products

PUBLICATIONS. DATA. SOFTWARE. +

Open Research Information and Analytics

- **Open REST APIs**

Graph search (<https://graph.openaire.eu/docs/apis/graph-api>)

Graph citations (<https://graph.openaire.eu/docs/apis/scholexplorer/api>)

- **Open datasets**

Zenodo.org (<https://graph.openaire.eu/docs/category/downloads>)

Beginners kit (<https://graph.openaire.eu/docs/downloads/beginners-kit>)

- **Data analytics environment**

Today: Google Cloud ([OpenAIRE AMKE](#) and [ORION-DBs](#), Kramer, Neylon et al.)

Tomorrow: Investigating move of data and analytics to the EOSC EU Node

- **Agentic AI**

MPC servers - Alien

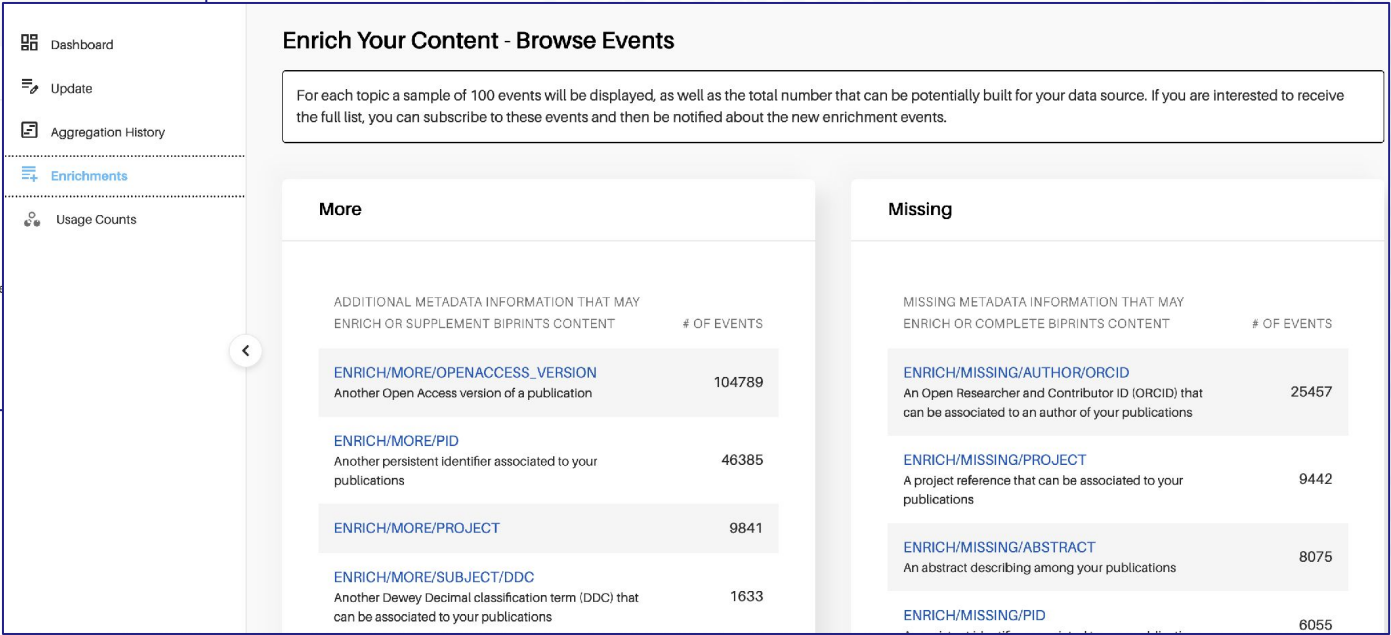
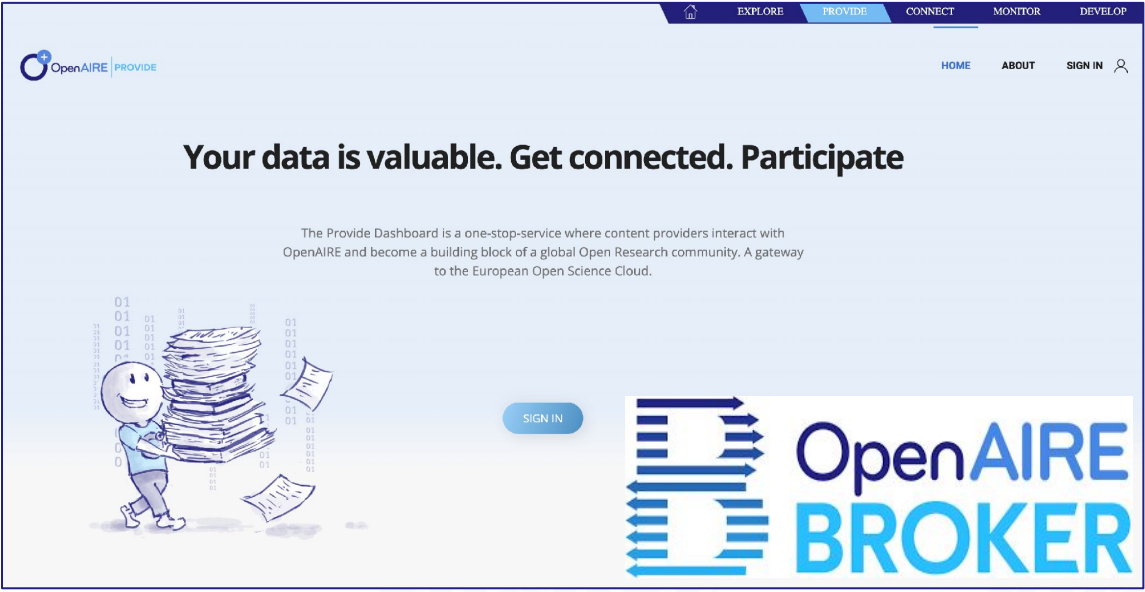
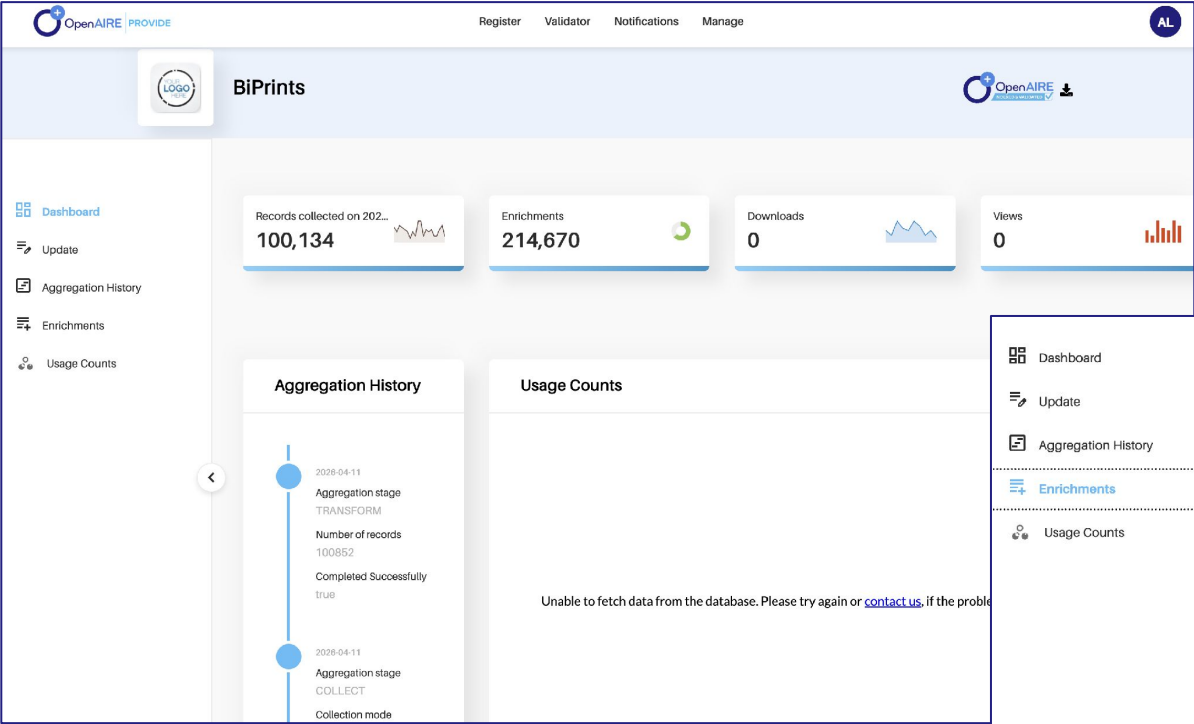


Want to know more: <https://graph.openaire.eu/docs>

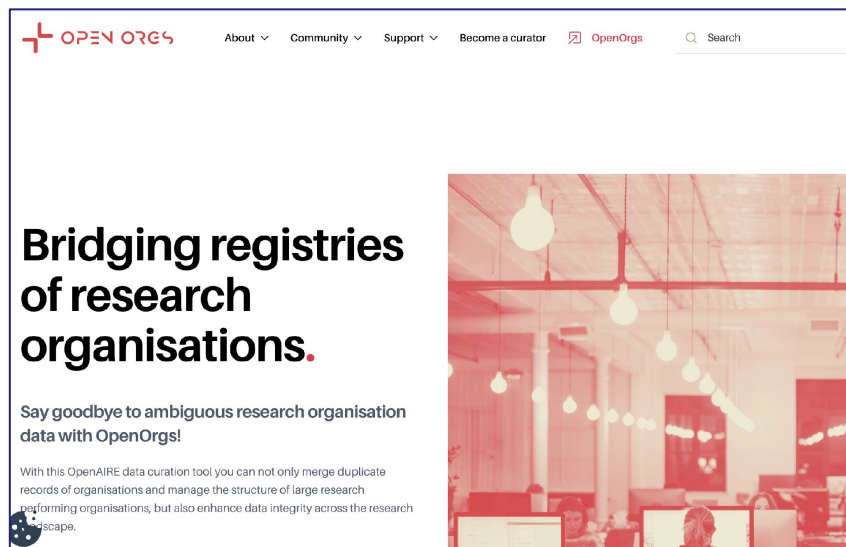
Virtuous exchange of research information

“High-quality research Information should persist in repositories and publisher archives (not in Knowledge Graphs)”

OpenAIRE PROVIDE and Broker



OpenOrgs – curating organizations (since 2024)



The image shows the OpenOrgs website header with navigation links: About, Community, Support, Become a curator, OpenOrgs, and a search bar. The main content area features the headline "Bridging registries of research organisations." and a sub-headline "Say goodbye to ambiguous research organisation data with OpenOrgs!". Below this, a paragraph states: "With this OpenAIRE data curation tool you can not only merge duplicate records of organisations and manage the structure of large research performing organisations, but also enhance data integrity across the research landscape." A red-tinted photograph of a modern office interior with hanging light bulbs is positioned to the right of the text.

OpenOrgs in numbers.

Discover key statistics and insights from our comprehensive dataset of organisations worldwide. Our dedicated team of 70+ curators from 20+ collaborating countries works tirelessly to ensure the data's accuracy and relevance, contributing to a rich and dynamic database.



The infographic consists of two parts. On the left, there are 70 red human icons arranged in four rows of 17, with the last row having only 6 icons. One icon in the third row is highlighted with a red circle and the text "this one might be you!". On the right, there is a map of Europe with 20 countries highlighted in red.

70+	20+
curators	collaborating countries

Types of organisations

From universities and research institutes to funding bodies, OpenOrgs organizes entities into categories that highlight their unique functions within the broader research and academic landscape.

Curated entities			
 2728 Archive	 28003 Company	 19870 Education	 8528 Facility
 5773 Government	 12474 Healthcare	 12571 Nonprofit	 7821 Other

Global curation process.

With over 102,000 organisations already meticulously curated, OpenOrgs is bridging continents, creating a comprehensive and accessible database that spans 181 countries and 56 territories and regions.

102k+ curated organisations	205k+ organisations left to curate	237 countries, territories and regions
---------------------------------------	--	--

The OpenAIRE Graph and Digital Sovereignty

The OpenAIRE Graph as “A reference implementation of the Barcelona Declaration principles”

Global coverage

Closed databases, open data collections, and a high-quality focus of the European research landscape

Beyond publications

publications, data, software, funding, grants, organisations (and more to come)

Transparent and Open

Data provenance, open data, open software

Trusted

Not just technical decisions, but design decisions driven by political choices and community demands

Infrastructure, not a service:

Services compete. Infrastructure connects.

Community governance

Provides full control on what and how should be tracked and monitored

Intertwined with Open Science roadmaps

Shaping priorities in the open, via the EC, its members, and participation to international fora

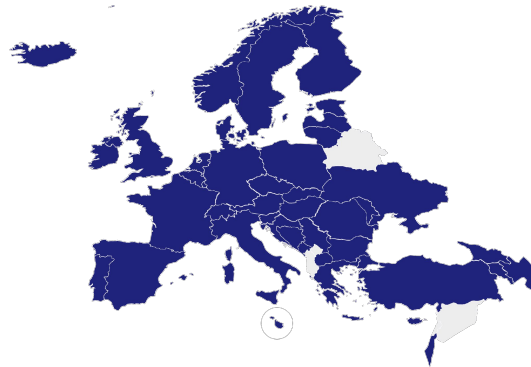
Towards community governed research information

01

Fostering full control on the research intelligence ecosystem

Without a public, governed knowledge base, public institution, funders, and communities outsource strategic decisions about science, innovation, and assessment. The OpenAIRE Graph is operated by the community, represented by OpenAIRE AMKE (no-profit legal entity, 2018), its members, Steering Committees, collaborating initiatives.

Global Collaborations: Latin America, Canada, Japan, Korea, China,...



62

members

36

countries

Common standards improve the network at its source

02

Global alignment reduces repair downstream.

OpenAIRE contributes to the co-design of guidelines, vocabularies, PID practices, help repositories and publishers share interoperable metadata. This lowers future curation costs and makes research information easier to exchange beyond OpenAIRE.



- *Scholix.org*
- *Scholarly Knowledge Graphs Interoperability Framework*
- *DMP data models*



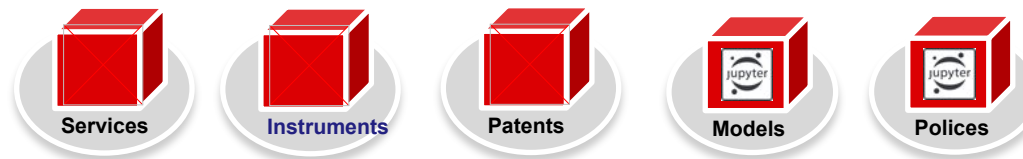
- *OpenAIRE guidelines for institutional repositories*
- *OpenAIRE guidelines for CRIS systems*
- *Platform integration*

Research assessment in the hands of the community

03

The community is to establish what should be monitored, assessed, and how.

The OpenAIRE Graph evolves to match the needs of the community, from entities, to links, and vocabularies. New entities, new data sources, new classifications are introduced to match the needs of monitoring and assessment.



Economy of scale of research information

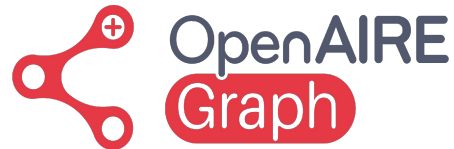
04 Curate in the Graph, without rebuilding another local research information collection

The OpenAIRE Graph absorbs the recurring cost of collecting, cleaning, deduplicating, linking and enriching research information. Consumers can start from open, ready-to-use data and invest in their questions, indicators and services instead of rebuilding the same foundation.

Every correction improves a shared public resource

05 One curation effort benefits every present and future user

A better affiliation, a resolved duplicate, a newly inferred project link or an improved classification enters the common Graph. The result does not remain inside one private dashboard or one institutional data pipeline.



Better information flows back to repositories

06 The Graph enriches the sources that sustain it.

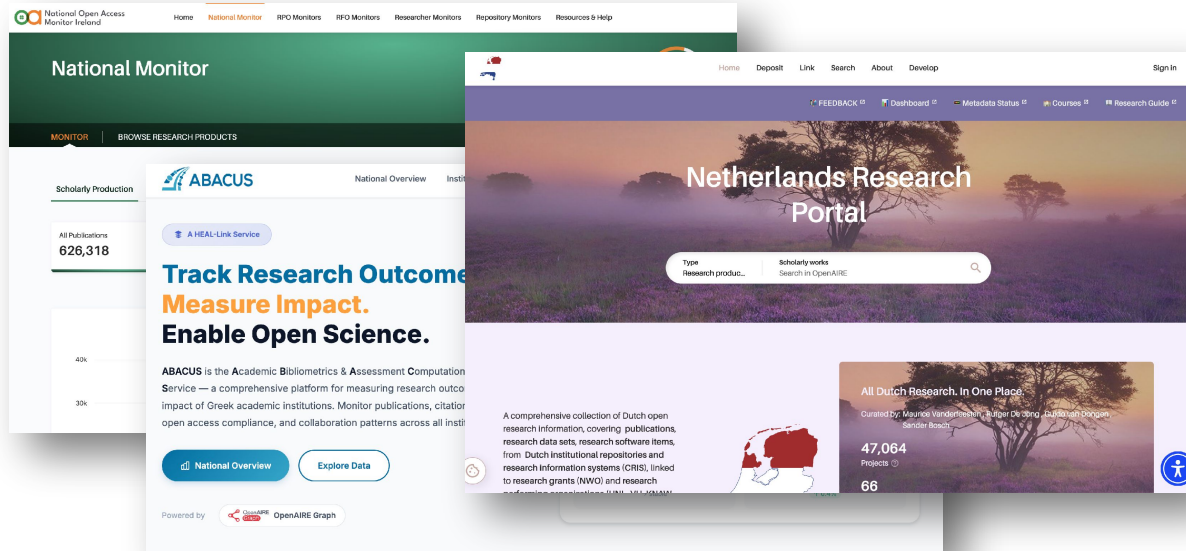
OpenAIRE PROVIDE and the Broker return links, corrections and enrichment opportunities to repositories. Authoritative information remains with the institutions that steward it, rather than living only in a central aggregation layer.



Open Research Intelligence

Real use-case examples

Institutional



On-demand dashboards providing timely, accurate, and well-documented indicators for research activities, outputs, and impact.

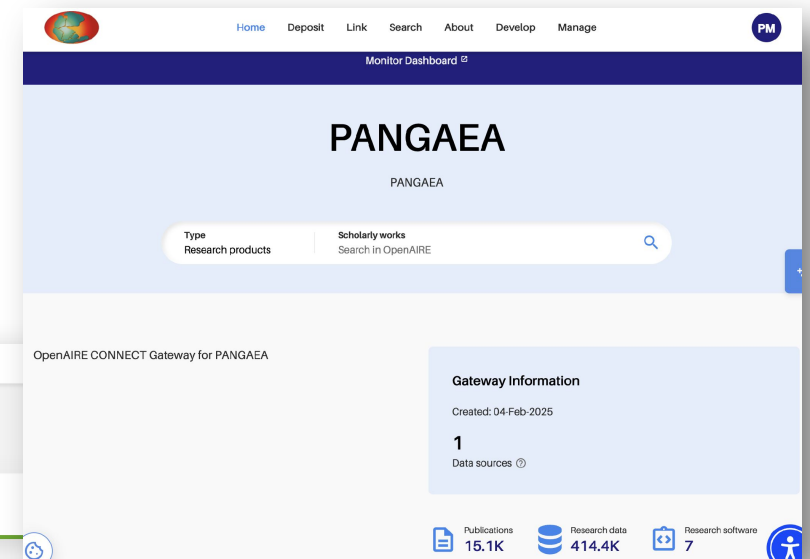
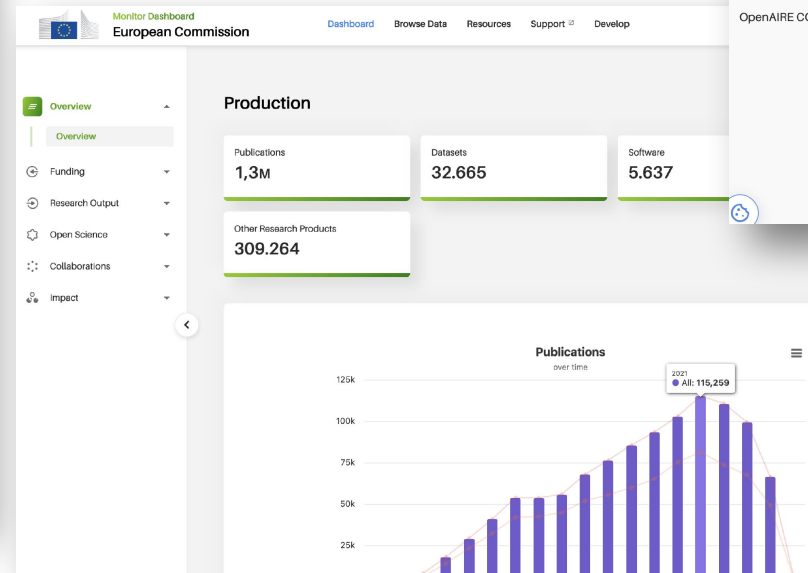
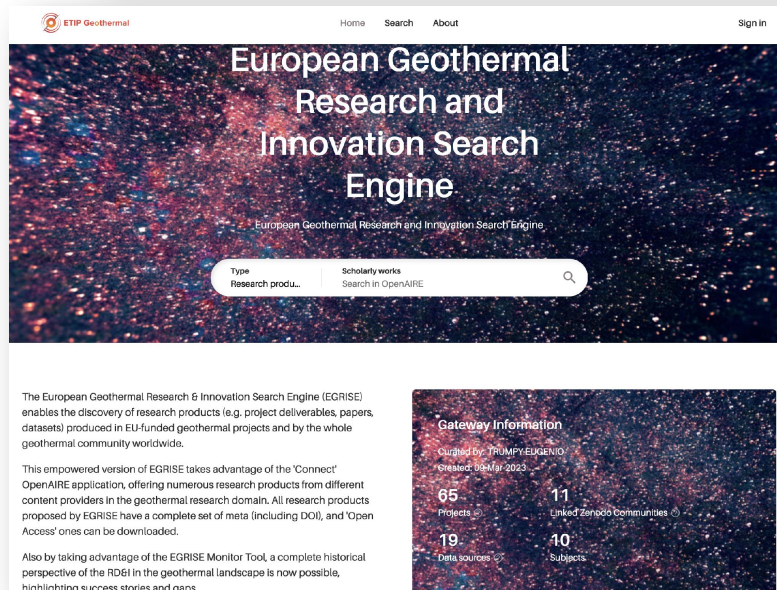
Designed for research organisations, funders, and communities.

Tailored to advance Open Science.

Used by

112	25	13	67	7
All Dashboards	Research Institutions	Funders	Research Initiatives	Publishers

Funders



Repositories & Publishers

Thematic

Agentic AI: OpenAIRE Graph MCP connector

Please
produc



I'll produce a comprehensive report on the University of Padua's research output over the last 5 years. Let me start by finding the relevant data.

Searched available tools

Let me start by finding

Search Organizations >

These are all departments

Used OpenAIRE integration

I found the main University of Padua

openorgs_____: :d6dd

projects involving the

Used OpenAIRE integration

Let me try different funding sources

Search Projects >

I see — the funder should be the University of Padua (Ricerca), which is the main research/university body funded by "MIUR". Let me query

Search Projects >

I found 8 MIUR/MUR-funded projects at University of Padua. The dates appear to have parsing issues (showing as "0029" instead of "2029" — a problem in OpenAIRE). Now I need to analyze the results and compute the Open Access ratio. Let me search for Open Access projects linked to the main projects.

Searched available tools >

Now I have the right tool to search research products. I found 8 projects from University of Padua, funded by MIUR/MUR, in the last 5 years. Let me check their open access status.

Let me first get the total count of research products linked to the main funder over the past 5 years:

Search Research Products >

Excellent — the total is 1,566 research products linked to the main funder from 2021 to 2025. Now I need to compute the Open Access status individually.

Search Research Products >

1,292 OPEN out of 1,566 total. Let me get each remaining project's status.

Used OpenAIRE integration >

So the access status breakdown is: OPEN=1292, CLOSED=143, EMBARGO=4, UNKNOWN=0. Total = 1561 (vs the 1566 — a rounding/categorization gap of 5). Let me now break down the Open Access distribution and Green OA.

Key data points collected:

UniPD + MIUR/MUR funded research products, 2021-2025:

- **Total: 1,566 research products** (1,544 publications, 13 datasets, 9 "other", 0 software)
- **Open Access (OPEN): 1,292 (82.5%)**
- **Closed: 143 (9.1%)**
- **Restricted: 122 (7.8%)**
- **Embargo: 4 (0.3%)**

OA color breakdown (within the 1,292 OPEN):

- Gold: 493
- Hybrid: 520
- Bronze: 36
- Diamond journals: 9
- Green OA: 1,120 (deposited in repositories — overlaps with above)

Year-by-year:

- 2021: 299/399 = 74.9%
- 2022: 233/302 = 77.2%
- 2023: 217/253 = 85.8%
- 2024: 267/300 = 89.0%
- 2025: 276/312 = 88.5%

Field of Science (FOS) breakdown:

- Natural sciences: 351
- Medical and health: 429



Claude



alien intelligence

Report as docx

Time to deliver: 12 minutes



3.2 Trend over the five-year window

The annual progression shows a clear upward trajectory. Two regimes are visible: a slower 2021-2022 baseline at roughly three-quarters OA, and an accelerated 2023-2025 plateau in the high 80s. The inflection corresponds in time to the publication of the National Plan for Open Science (mid-2022) and to the activation of UniPd's expanded transformative agreements through the CRUI-CARE consortium.

Year	Total outputs	OA outputs	OA ratio	Δ vs prior year
2021	399	299	74.9%	—
2022	302	233	77.2%	+2.3 pp
2023	253	217	85.8%	+8.6 pp
2024	300	267	89.0%	+3.2 pp
2025	312	276	88.5%	-0.5 pp
5-year total	1,566	1,292	82.5%	

Reading the table: The total-output dip in 2022-2023 is partly an artefact of indexing latency (older funder linkages keep arriving in OpenAIRE for several months after publication) and partly real, reflecting the transition between PRIN/FARE funding cycles. The OA ratio is the more reliable indicator: it strengthens consistently and stabilises near 89% from 2024 onward.

3.3 Composition of the open share

Within the 1,292 openly accessible outputs, OpenAIRE's colour taxonomy distinguishes the publication route taken. Note that categories overlap — most green-deposited articles also carry a publisher-side colour (gold, hybrid, or bronze).

Open Access route	Outputs	Share of OA total	Description
Green (repository deposit)	1,120	86.7%	Self-archived in IRIS-PRA or another OpenAIRE-compliant repository
Hybrid	520	40.2%	OA article in a subscription journal (often via APC or transformative agreement)
Gold	493	38.2%	Article in a fully OA journal
Bronze	36	2.8%	Free to read on the publisher site without an open licence
Diamond	9	0.7%	OA journal with no APC, fully community-supported



Thank you!

Paolo Manghi

paolo.manghi@openaire.eu

Images licence
<https://www.istockphoto.com/legal/license-agreement>