

 argos

 OpenAIRE
ResearchGraph

FAIR ma-DMPs & Science Knowledge Graphs

The OpenAIRE services

Natalia Manola

CEO, OpenAIRE

Acknowledgements



Elli Papadopoulou
Athena Research & Innovation Center



Paolo Manghi
CNR & OpenAIRE CTO



George Kakalettris
Cite SA

DATA MANAGEMENT PLANS - TODAY

Policy

- Public DMPs
- Rich content
- Support

Technical

- Search & retrieve
- Interoperability
- Semantics

Adoption

- FAIR practices
- Disciplines / Communities
- Data sharing

- Not Open Access
- Granularity of content
- Unclear about the process, where to start, how to get input

- Resource type
- Reused datasets
- Qualified references

- Reproducibility
- Domain Data Protocols
- Input & Output Data

SCIENCE KNOWLEDGE GRAPHS - TODAY

- **Quite a few exist already** - different use, formats, maturity
 - Scholarly Communication (OpenAIRE Graph, OpenAlex, +commercial ones)
 - National CRIS Systems (PT, FI, RS, EL, ...) - focus on monitoring
 - Domain disciplines ones (EBRAINS, ENVRI, ECRIN, ...) - focus on discovery
- **Challenges**
 - Most looking into publications
 - Some semantics in place, still to come
 - Partial coverage
 - No standards for access or exchange of information

The logo for Scilake, featuring the word "Scilake" in a bold, dark blue sans-serif font. The letter "i" is replaced by a red dot, and the letter "l" is replaced by a teal dot. A thin horizontal line connects the red and teal dots above the "i" and "l".

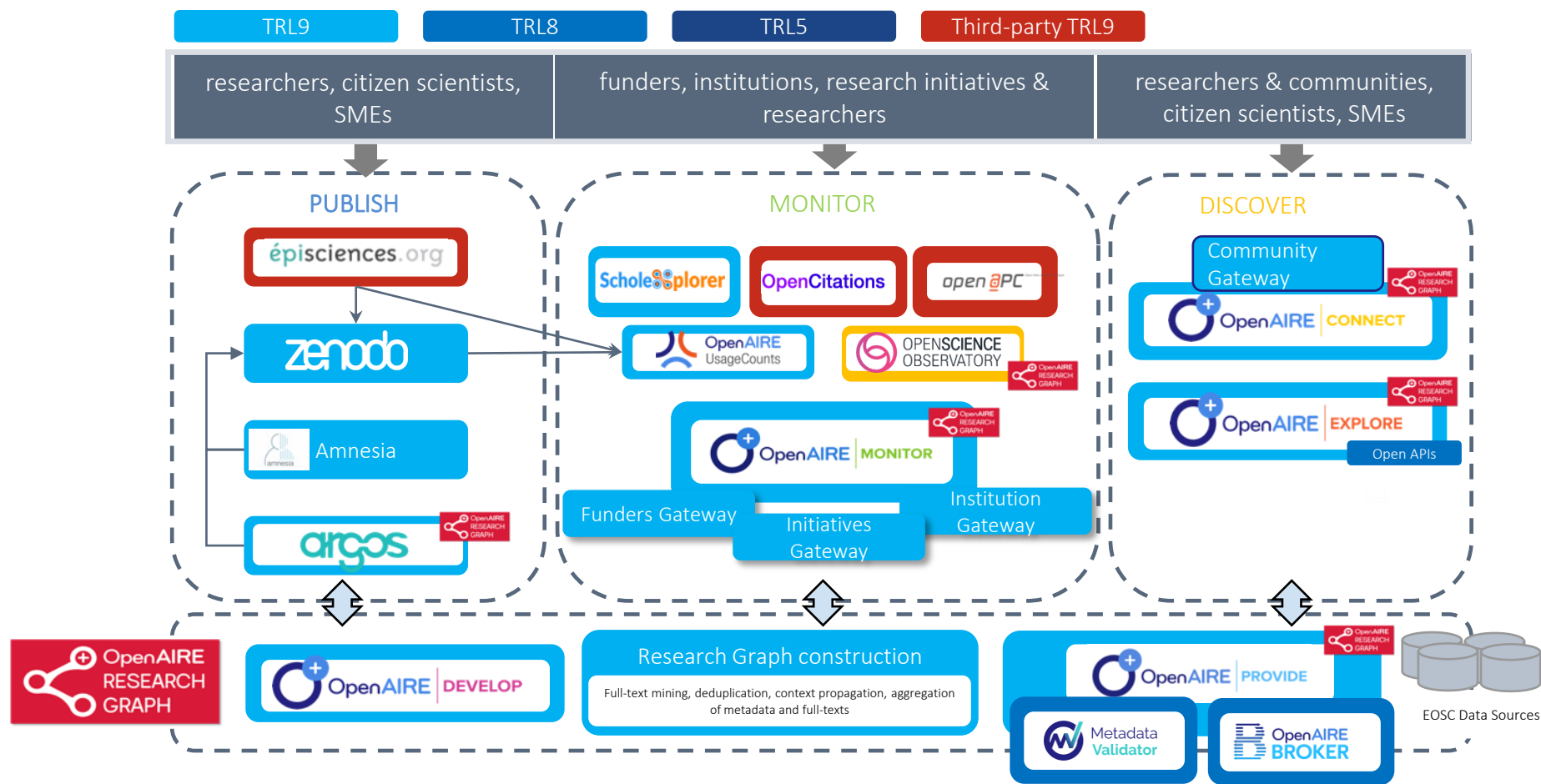
Scilake

Horizon Europe project starting in Jan 2023

The OpenAIRE view

end-to-end solutions

OPENAIRE SERVICES – ALL INTERCONNECTED

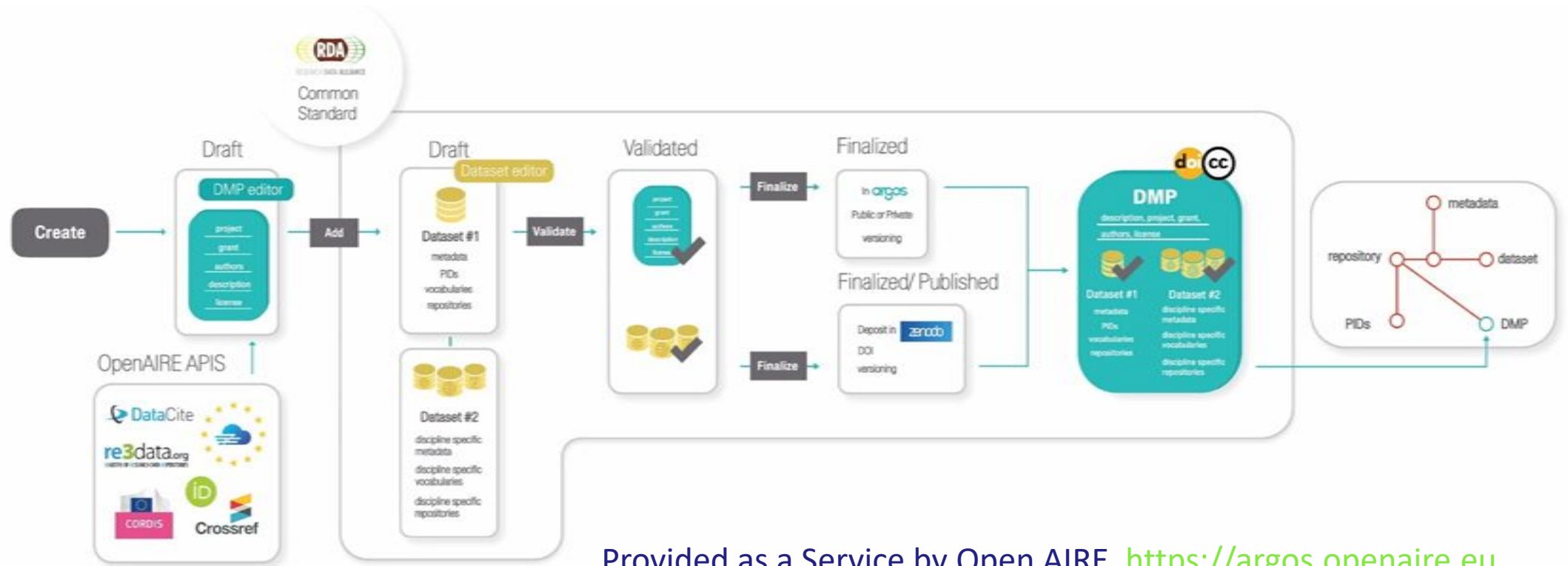


DATA MANAGEMENT PLANS

– IN / WITH ARGOS

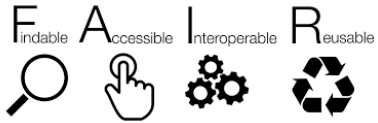
argos – A GLIMPSE

ARGOS is an open source, configurable and extensible tool for **planning Research Data Management (RDM)** activities according to **Open Access & FAIR data policies**.



Provided as a Service by Open AIRE <https://argos.openaire.eu>

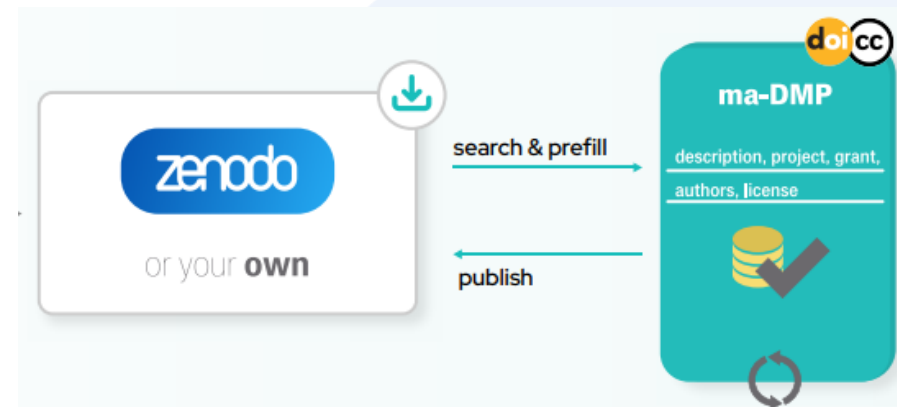
argos – OUTPUTS & PROCESSES



- **Findable:** OpenAIRE EXPLORE
- **Accessible:** PIDs (ORCIDs & DOIs)
- **Interoperable:** RDA DMP Common Standard
- **Reusable:** Licenses
- **Versioned:** history/provenance
- **Published and preserved** in Zenodo

• Automated

- **Writing:** inferred content and **allocation** in Templates as answers
- **Searching:** Published Dataset / Software metadata
- **Publication:** DMP metadata and files in Zenodo



UNLOCKS NEW POTENTIALS – NOT JUST A TOOL!



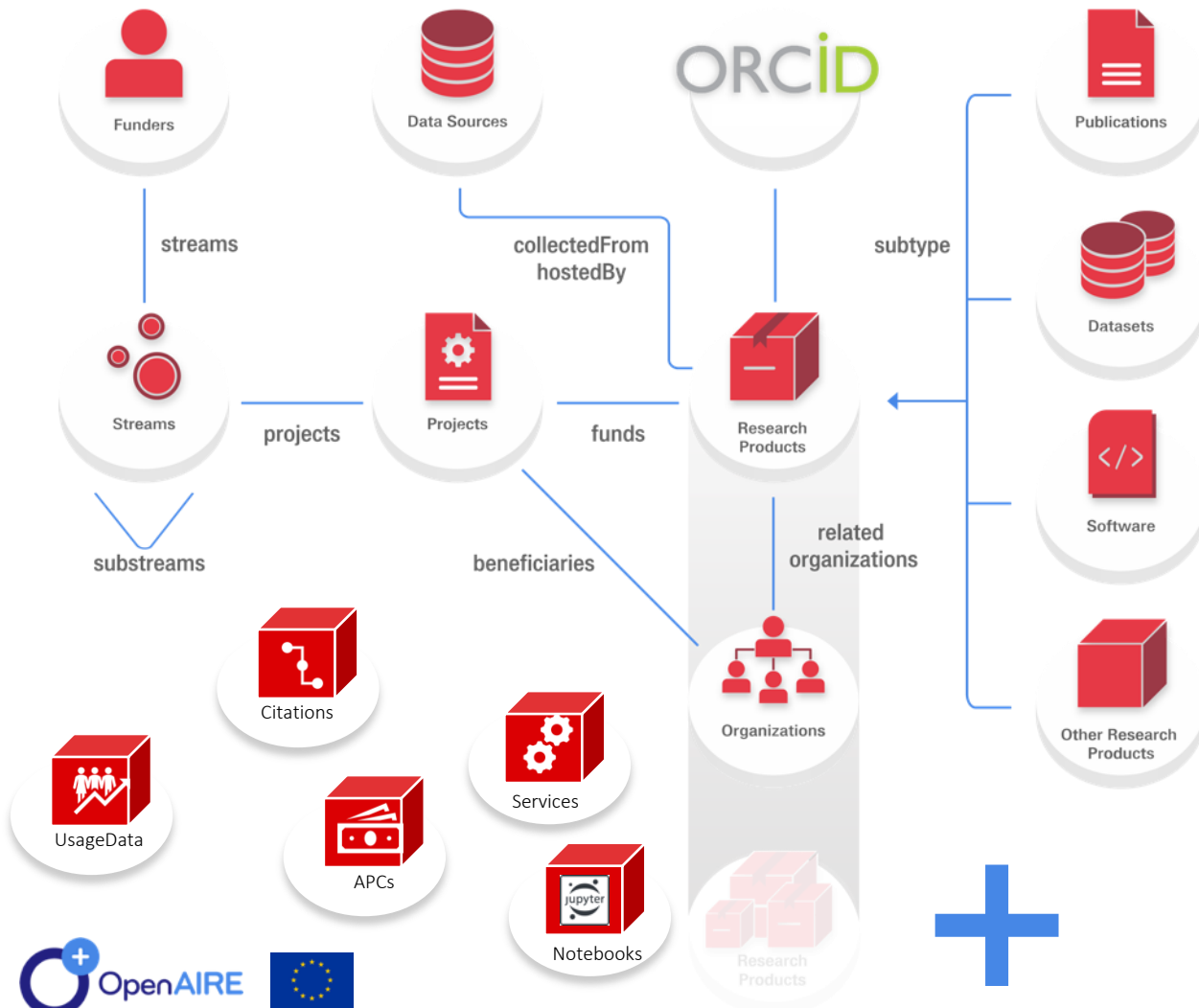
- **For all outputs**
 - Data, Software, Workflows, etc.
- **For all disciplines**
- **Connects data services and workflows**
 - Validation, Automation, Easy DMP publication with a click

SCIENTIFIC KNOWLEDGE GRAPH

– THE OPENAIRE GRAPH

OPENAIRE **GRAPH** – KEY ASSET IN OPEN SCIENCE INFRASTRUCTURE

INCLUSIVE, TRANSPARENT, REPRODUCIBLE

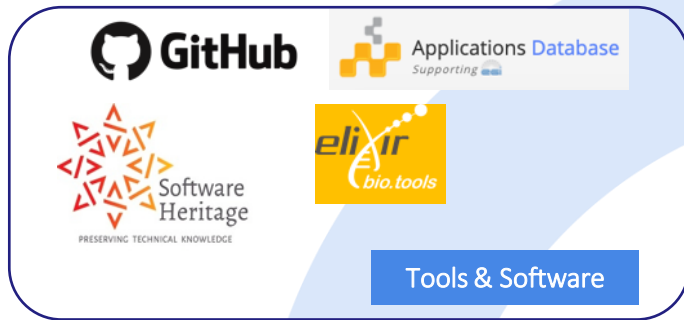


- Based on **Open Science principles**: open & authoritative sources, open APIs, well established **metrics**
- **Global** coverage
- Enriched by **AI-processes** - citation, reproducibility, FOS, SDGs, domain specific
- Moving towards **smart discovery**
- High **relevance** to the research community

A comprehensive and open dataset of research information covering 145m publications, 18m research data, 305k research software items, from 110k data sources, linked to 3m grants and 178k organizations.
All linked together through citations and semantics.



OPENAIRE GRAPH DATA SOURCES



Tools & Software

Logos: GitHub, Applications Database, Software Heritage, elixir bio.tools



Publishers

Logos: frontiers, PLOS, Springer, F1000 Faculty of 1000



35 Funding databases

Logos: CORDIS, NSF, DFG Deutsche Forschungsgemeinschaft, NIH



Research Graphs

Logos: unpaywall, OpenCitations, Microsoft



OpenAIRE RESEARCH GRAPH



Registries

Logos: DataCite, re3data.org, ROR, ORCID, OpenDOAR, FAIRsharing.org standards, databases, policies, Crossref



Aggregators

Logos: LA Referencia Red de repositorios de acceso abierto a la ciencia, SciELO, DOAJ DIRECTORY OF OPEN ACCESS JOURNALS



Repositories / Pre-prints

Logos: Cornell University Library, arXiv.org, RePEc, HAL archives-ouvertes.fr, PubMed, zenodo Research. Share.



RI Catalogues

Logos: DARIAH-EU, elixir, EPOS EUROPEAN PLATE OBSERVING SYSTEM, ECRIN EUROPEAN CLINICAL RESEARCH INFRASTRUCTURE NETWORK, EOSC-Life, egi

DATA WRANGLING CHALLENGES



Heterogeneity

Standards, crosswalks, community and cross-community view

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



Quality and 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 (Scientific Knowledge Graphs)

Open Science publishing workflows, RDA
OSG interoperability

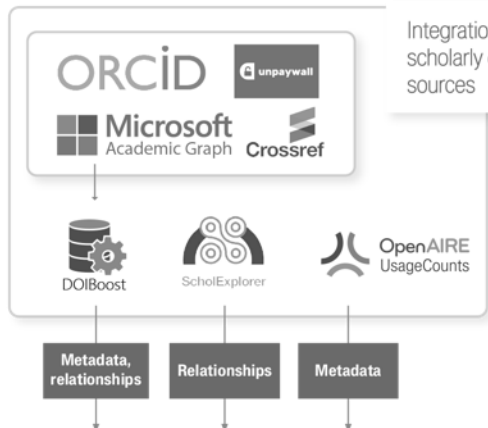
OPENAIRE GRAPH: THE SUPPLY CHAIN

1. Combining data from the global research ecosystem > 10K sources

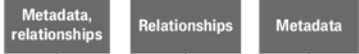


Different records representing the same entity (results or organization) are merged in one

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



Integration of additional scholarly communication sources



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

Metadata records corresponding

2. A big AI-driven data infrastructure to extract information (meaning)

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

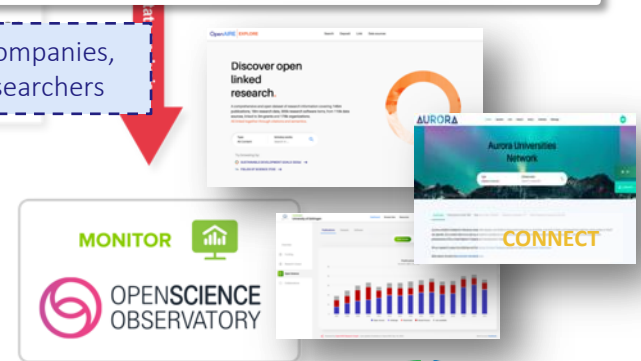
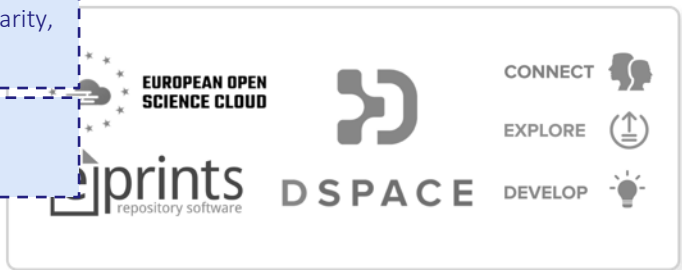
Propagation of properties via relationships

Inference of new properties and links via full-text mining

Inference of new properties and links from meta-data

3. Community to use the data

Third-party services: repositories, companies, SMEs, EOSC, Participant Portal, researchers



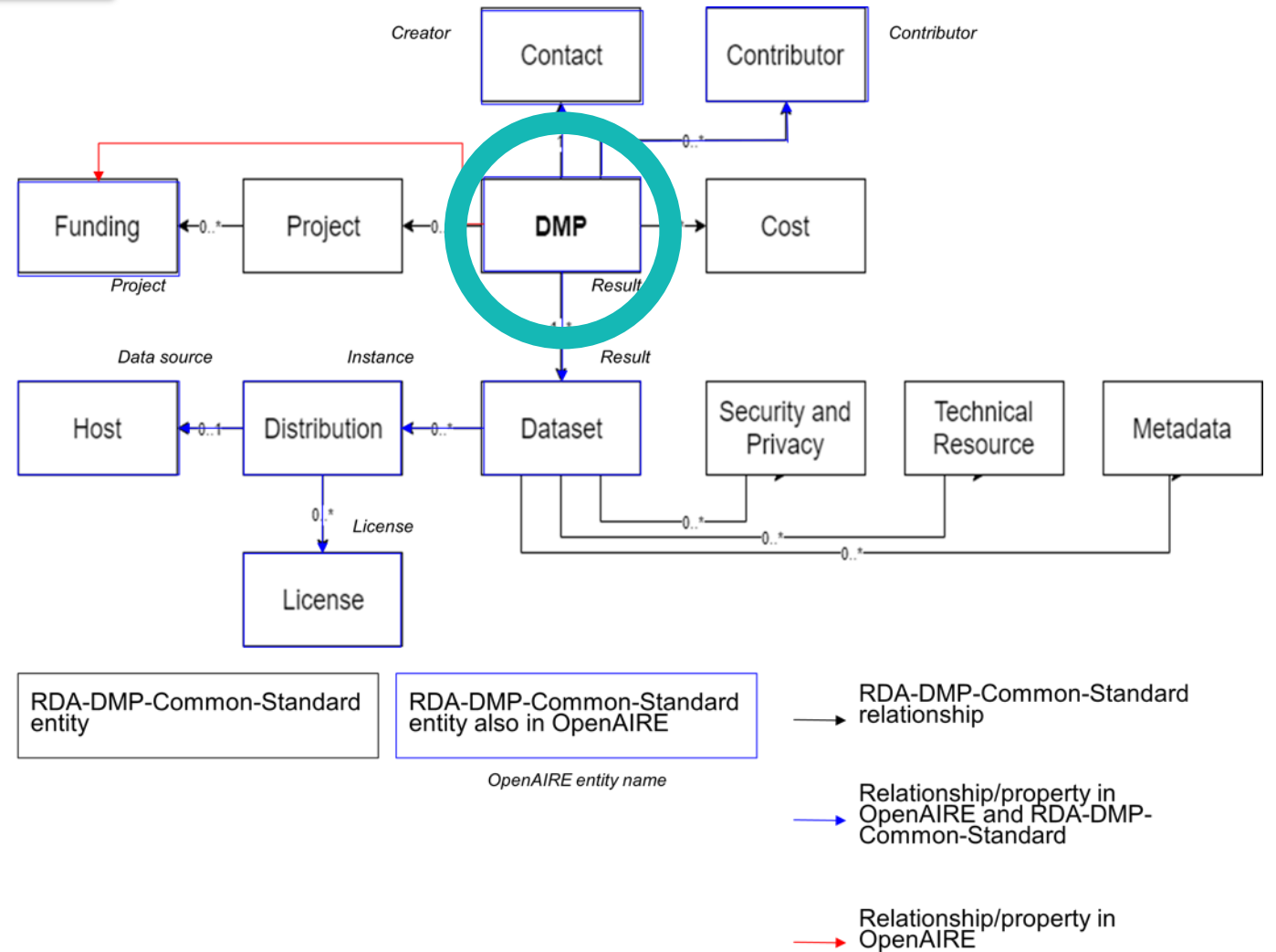


INTEGRATIONS IN PRACTICE

1.Manghi, Paolo, Bardi, Alessia, Atzori, Claudio, Baglioni, Miriam, Manola, Natalia, Schirrwagen, Jochen, & Principe, Pedro. (2019, April 17). The OpenAIRE Research Graph Data Model (Version 1.3). Zenodo. <http://doi.org/10.5281/zenodo.2643199>

MA-DMPS & OPENAIRE GRAPH

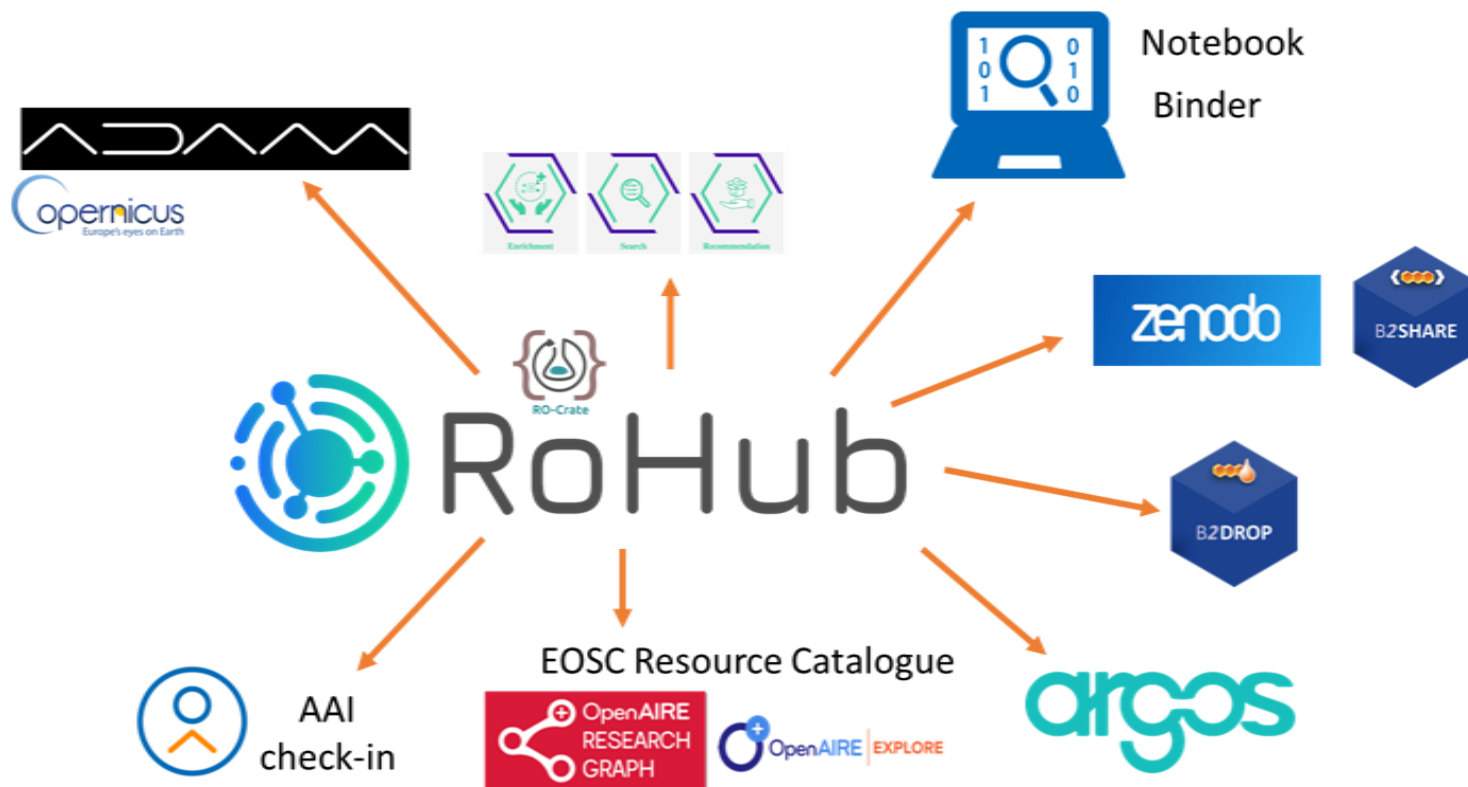
First integration



INTEGRATION WITH RO-HUB



<https://reliance.rohub.org/>



NEW IN 2023: RSPACE integrations for prefilling, linking and publishing information between the platforms

NATIONAL INSTANCES

The screenshot shows the Argos DMP interface. On the left is a navigation menu with items: Home, My DMPs, My Datasets, Public DMPs, Public Dataset Desc., DMP Templates, and Dataset Templates. The main content area features a modal window with the text: "A DMP in Argos consists of key information about research, such as purpose, objectives and researchers involved, but also about documentation of research datasets, namely **Datasets**, that highlight the steps followed and the means used across data management activities." Below this text is a yellow "Add Dataset" button and a cartoon character sitting at a laptop. At the bottom of the main area, there is a "Latest Activity" section with filters for "All (1)", "Drafts (0)", "DMPs (1)", and "Datasets (0)".



Poznańskie Centrum Superkomputerowo-Sieciowe
afiliowane przy Instytucie Chemii Bioorganicznej PAN

Connected to RO-Hub

<https://mitraria.man.poznan.pl/home>



Connected to Dataverse
institutional repositories



Change from other software: migration
mechanism for DMPs, users & their
connections

Integration with Dataverse repositories



in the pipeline

THANKS



Emails

- elli.p@athenarc.gr
- g.kakaletris@cite.gr
- Paolo.manghi@openaire.eu