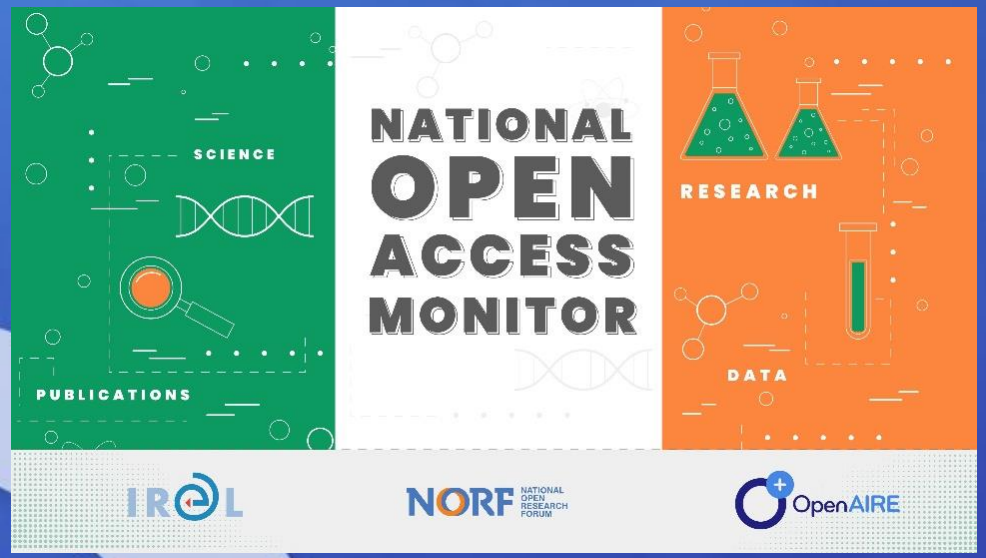
@openaire_eu

OpenAIRE
SCIENCE.SET FREE.



National Open Access Monitor, Ireland

Paving the road to 100% Open Access

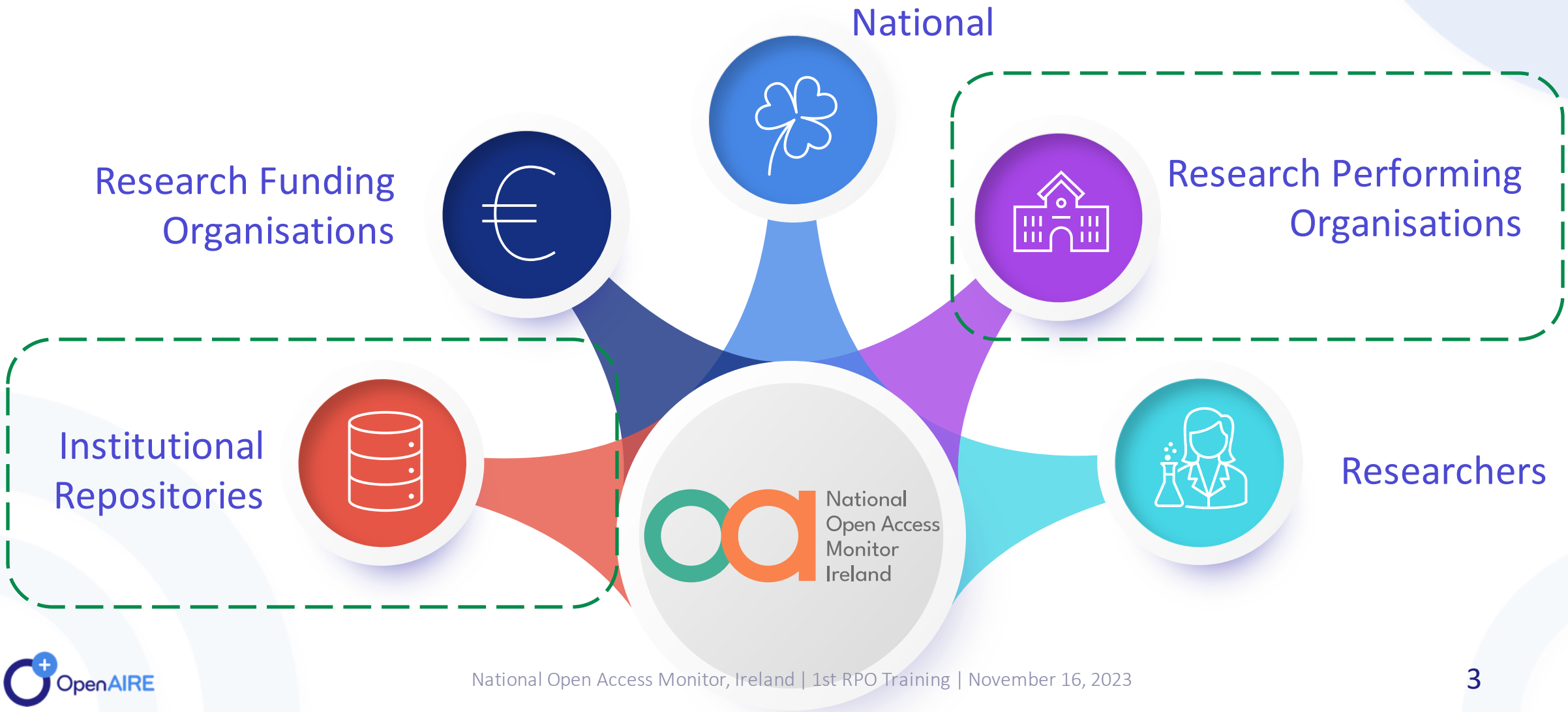
RPO Training

Leonidas Pispiringas, Data Scientist & Scholarly Communication Technical Expert

OUTLINE

- **Overview**
- **OpenAIRE Graph & Data Quality**
- **Linking functionality & Deduplication via OpenORGS**
- **Interoperability Guidelines & OpenAIRE PROVIDE**
- **Next Steps**

THE MONITOR – 5 DASHBOARDS



OBJECTIVES

RPOs

- **Why**
  - Establish national-level visibility for institutional research.
  - Compare OA performance against peer institutions.
  - Ensure compliance with Open Access mandates.
  - Uncover weak spots and insights to guide evidence-based decisions.

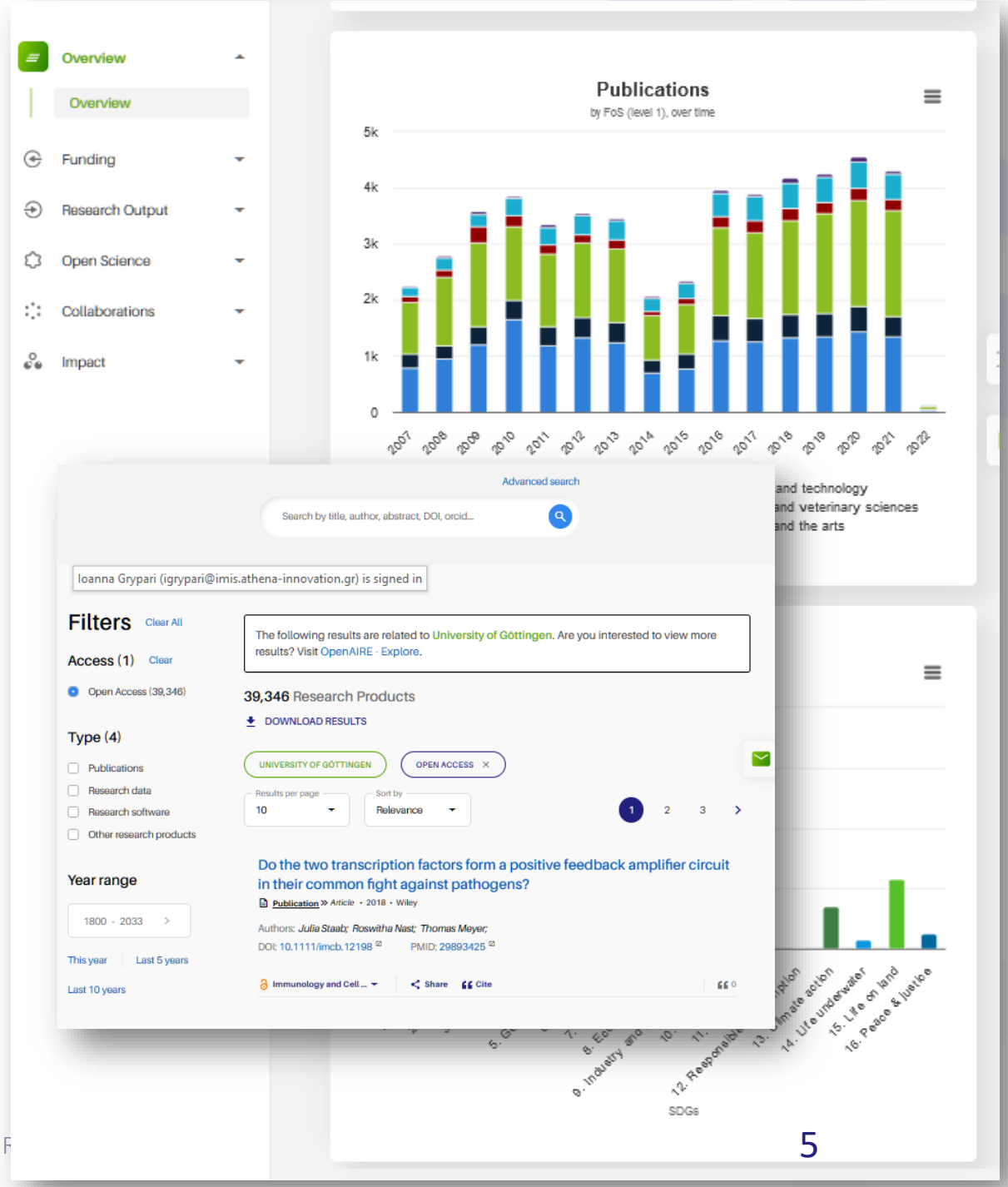
RPO Dashboards

- Comprehensive insights into each Irish RPO's activities and OS performance

- **Indicators View**
  - Highlighting **OA uptake** indicators, enables **cross RPO** benchmarking
  - Offers detailed **breakdowns** and **filtering**

- **Research Outputs View**
  - Facilitate **browsing**, standard and advanced **searching**, and **filtering**

- **RPO Managers**
  - Access to the **Sandbox** to verify data quality (e.g., if a new repository is harvested, deduplication in OpenORGS, etc.)
  - Access to **OpenORGS** tool to deduplicate organization name (RPOs)



Repository Dashboard

OBJECTIVES

Institutional Repository Managers

- **Why**
  - Ensure OA publications from their system are accurately reflected.
  - Facilitate the comprehensive coverage of the National Open Access Monitor of Ireland.

If registered with OpenAIRE PROVIDE

- Elevate repository records with precise validation and enrichment.
- Gain insights from detailed usage statistics.
- Simplify open access compliance, promoting effortless content dissemination.

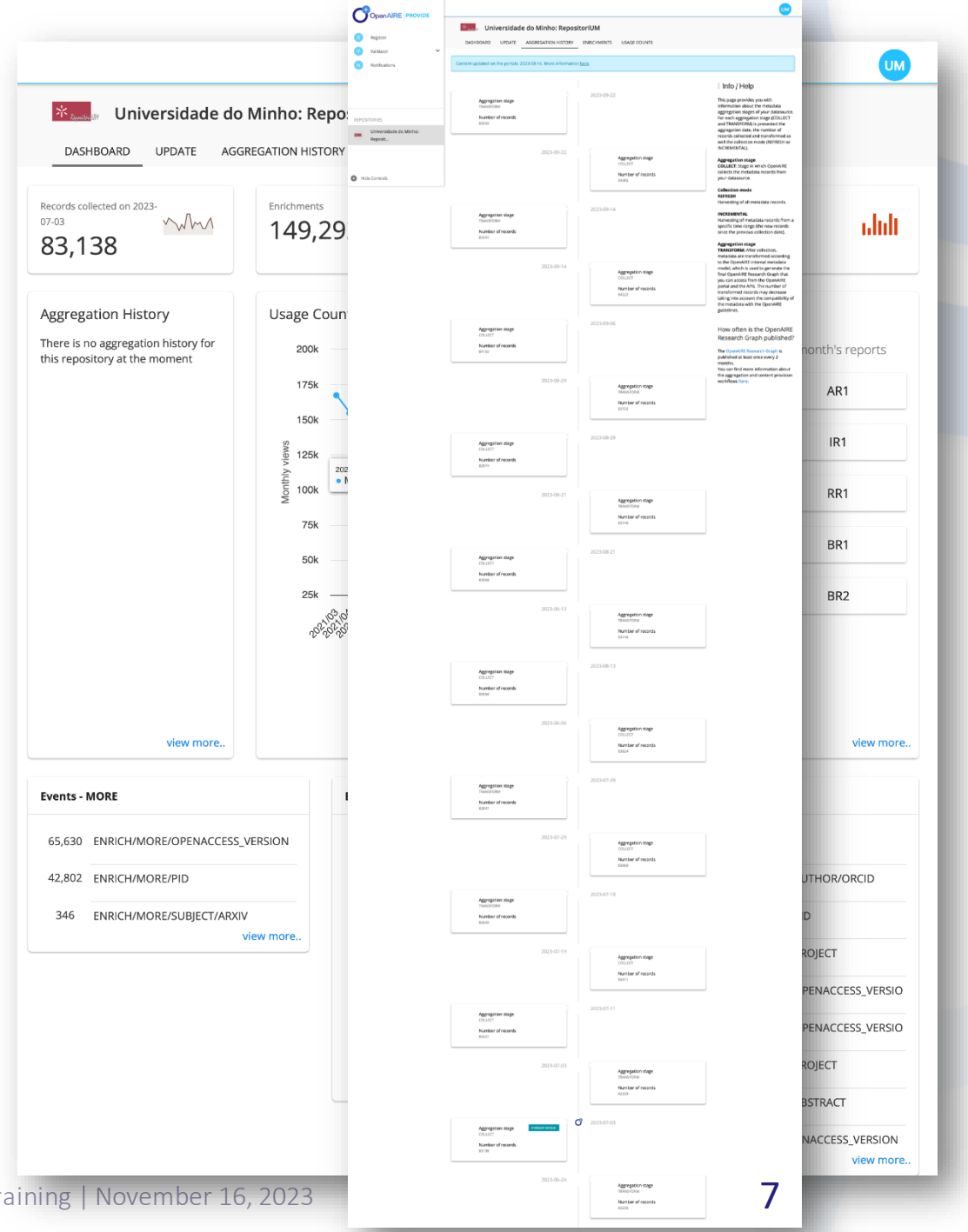
- Metrics and indicators on research output
- Detailed **breakdowns** and **filtering**
- B**rowsing**, standard and advanced **searching**, and **filtering**

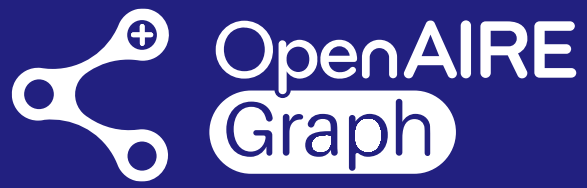
Full Functionality for Repositories registered in OpenAIRE

- Access to OpenAIRE PROVIDE Service dashboard
  - Register, validate (including FAIRness checks), and enrich records (via OpenAIRE BROKER)

Optional

  - Pre-requisite to align with the OpenAIRE Guidelines
  - Enrich their records' metadata from other sources harvested by OpenAIRE
- Usage Counts Service activation
  - Enrichment of the dashboard with usage reports for views and downloads data





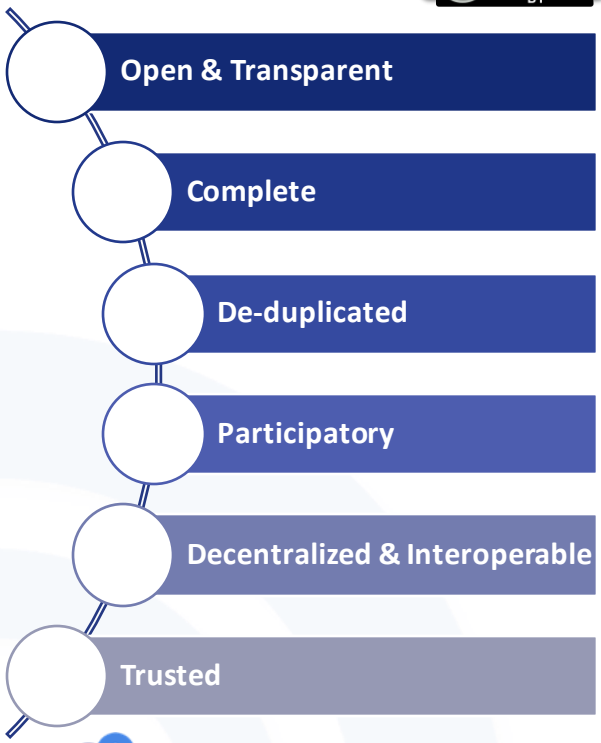
& Data Quality

THE OPENAIRE GRAPH

OpenAIRE Graph

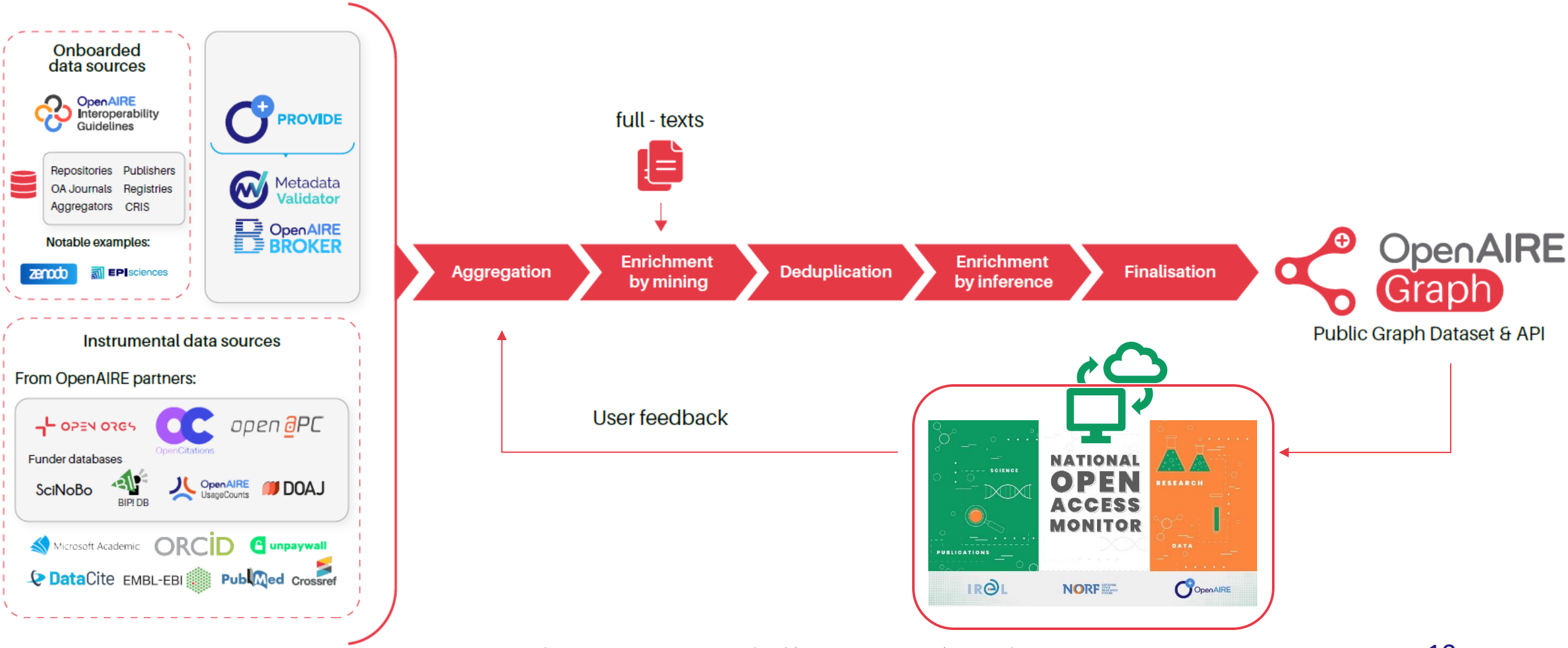
graph.openaire.eu

CC BY

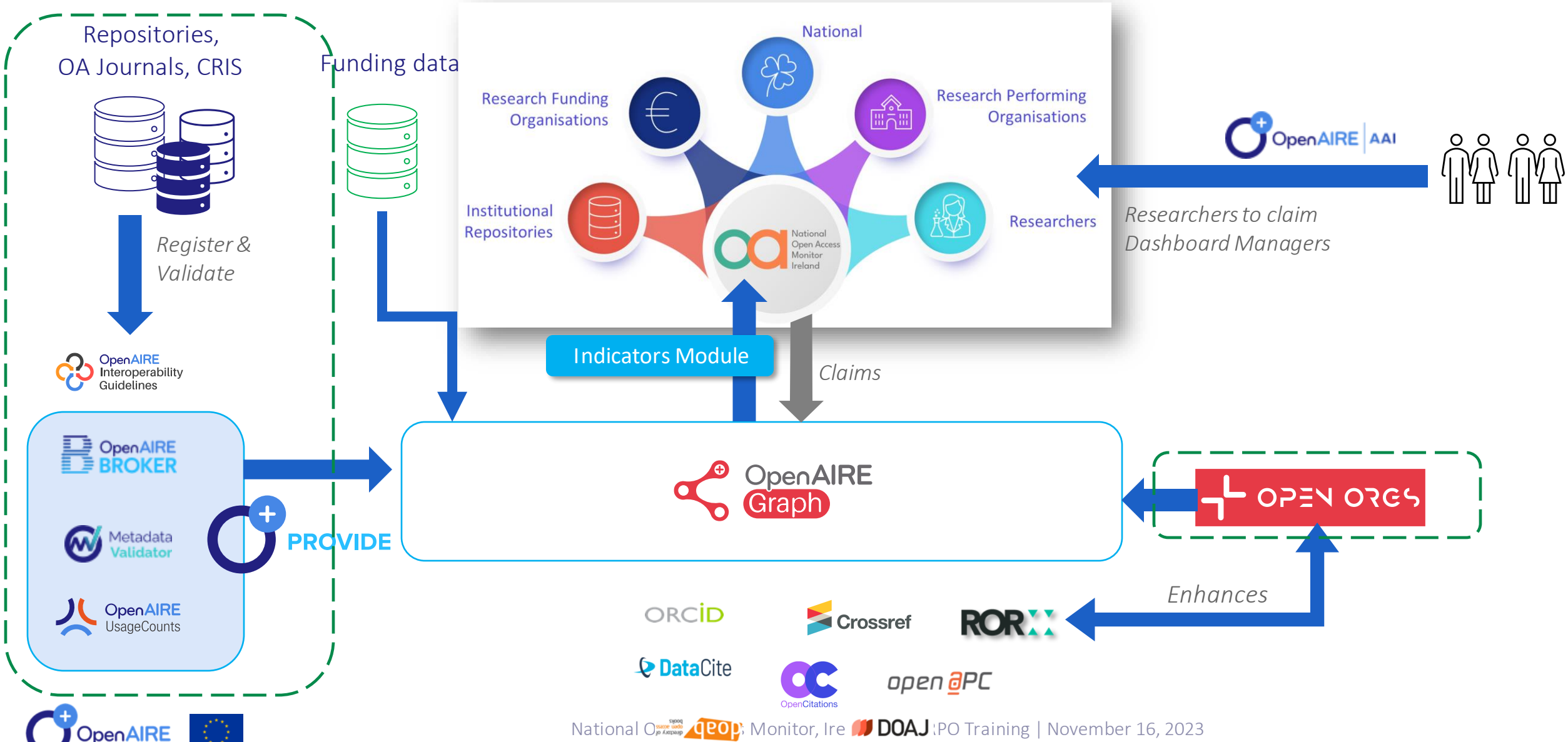


- **Scientific Knowledge Graph**
  - ***A collection of metadata describing objects in the research lifecycle and relationships among them***
  - Timely and comprehensive coverage of research outputs
  - 129K data sources, 3M projects, and 235M research outputs (publications, software, datasets, other research products)
  - Near monthly updates
- **Precision, depth & processing**
  - Rigorous cleaning, deduplication, and enrichment for optimal accuracy
  - Enriched metadata links: Research results to projects, author affiliations, and classifications (FoS, SDG)
  - Techniques for duplicate identification in addition to OpenAIRE's curation tools
- **Robustness & openness**
  - Maintenance, load balancing, backups, overseen by the OpenAIRE technology centre
  - Open data & transparent methodologies

GRAPH PIPELINE



THE ARCHITECTURE



COVERAGE: SOURCES CONTRIBUTING TO THE GRAPH

GitHub
Applications Database Supporting
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 graphs**

OpenAIRE Graph

DataCite
re3data.org
ROR
ORCID
Crossref
OpenDOAR
FAIRsharing.org standards, databases, policies
... 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
... and more **thematic and institutional repositories**

DARIAH-EU
elixir
EPOS
EOSC-Life
ECRIN
EGI
... and more **einfra and RI sources**

DATA QUALITY

- **Deduplication:** OpenAIRE Graph merges duplicate records of the same scholarly work.
- **Enrichment through TDM:** affiliation, citation, and concept extraction with document classification (FoS, SDG) and similarity assessment
- **Cleaning:** Independent, continuous aggregation processes utilize vocabularies to harmonize diverse data source records, ensuring consistent and accurate bibliographic records.
- **Disambiguation:** Additional disambiguation of Journals, Publishers and Licenses
- More @ https://graph.openaire.eu/docs/graph-production-workflow/

DATA QUALITY

Additional Data Quality Enhancements for RPOs & Repositories

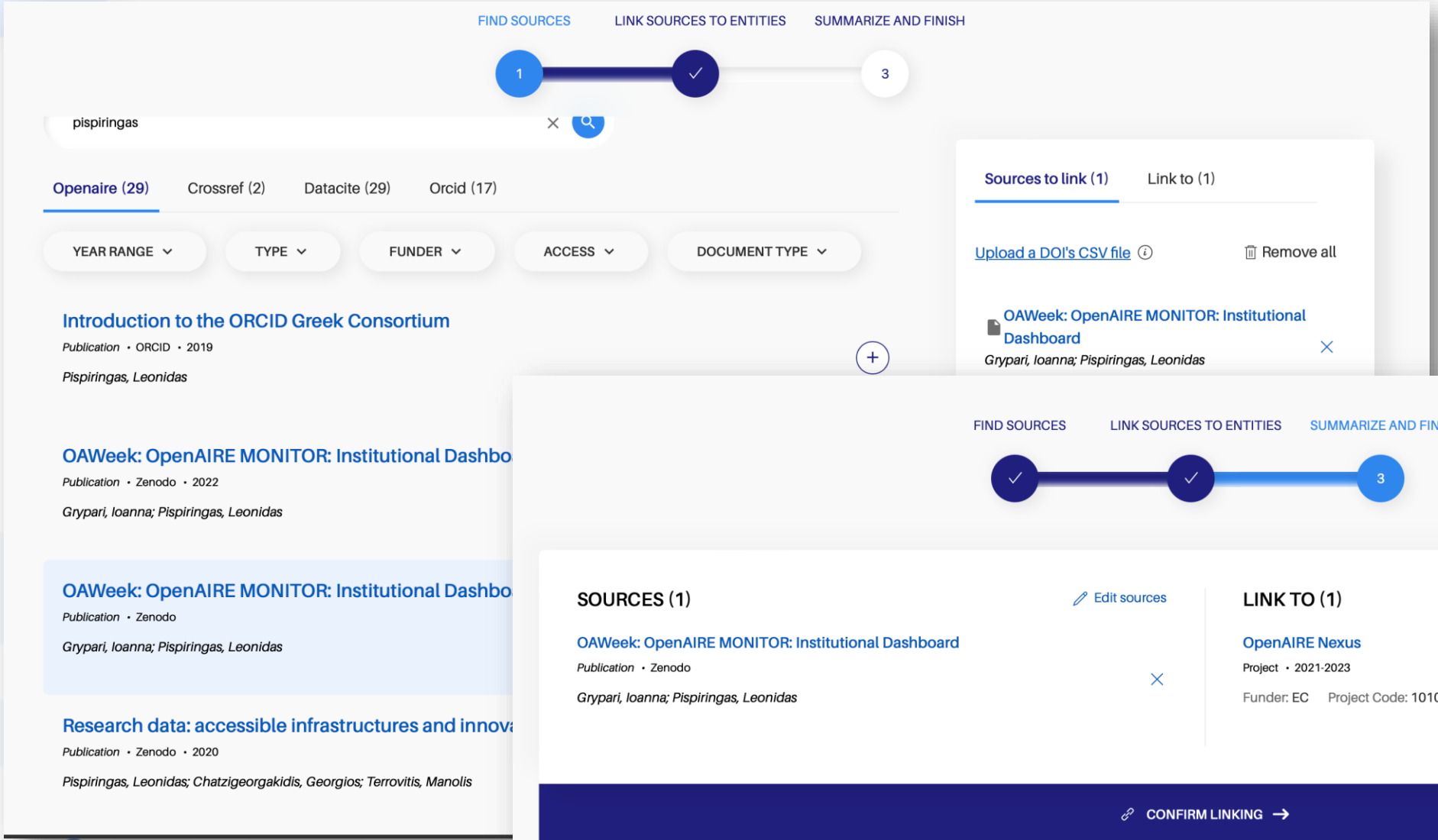
- **All**: via linking functionality
- **RPOs:** Deduplication via OpenORGS
- **Repositories:** Follow OpenAIRE Guidelines & Register with OpenAIRE Provide

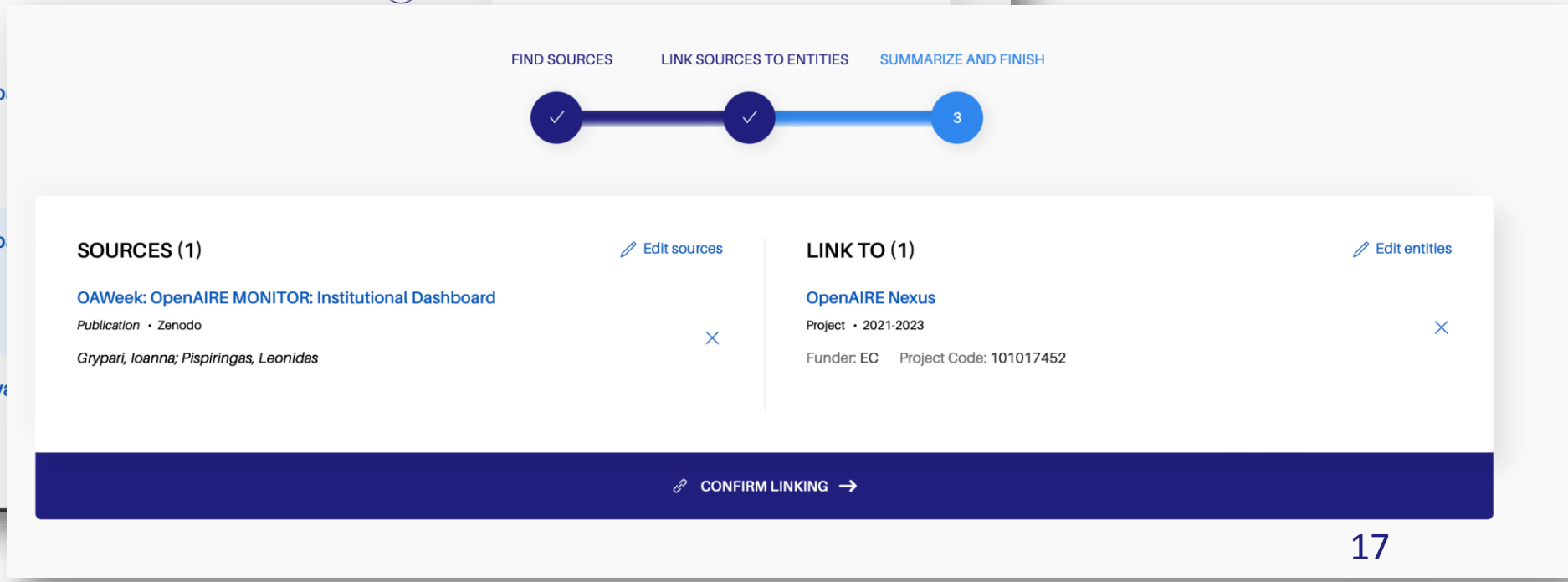
LINKING & OPENORGS

LINKING FUNCTIONALITY

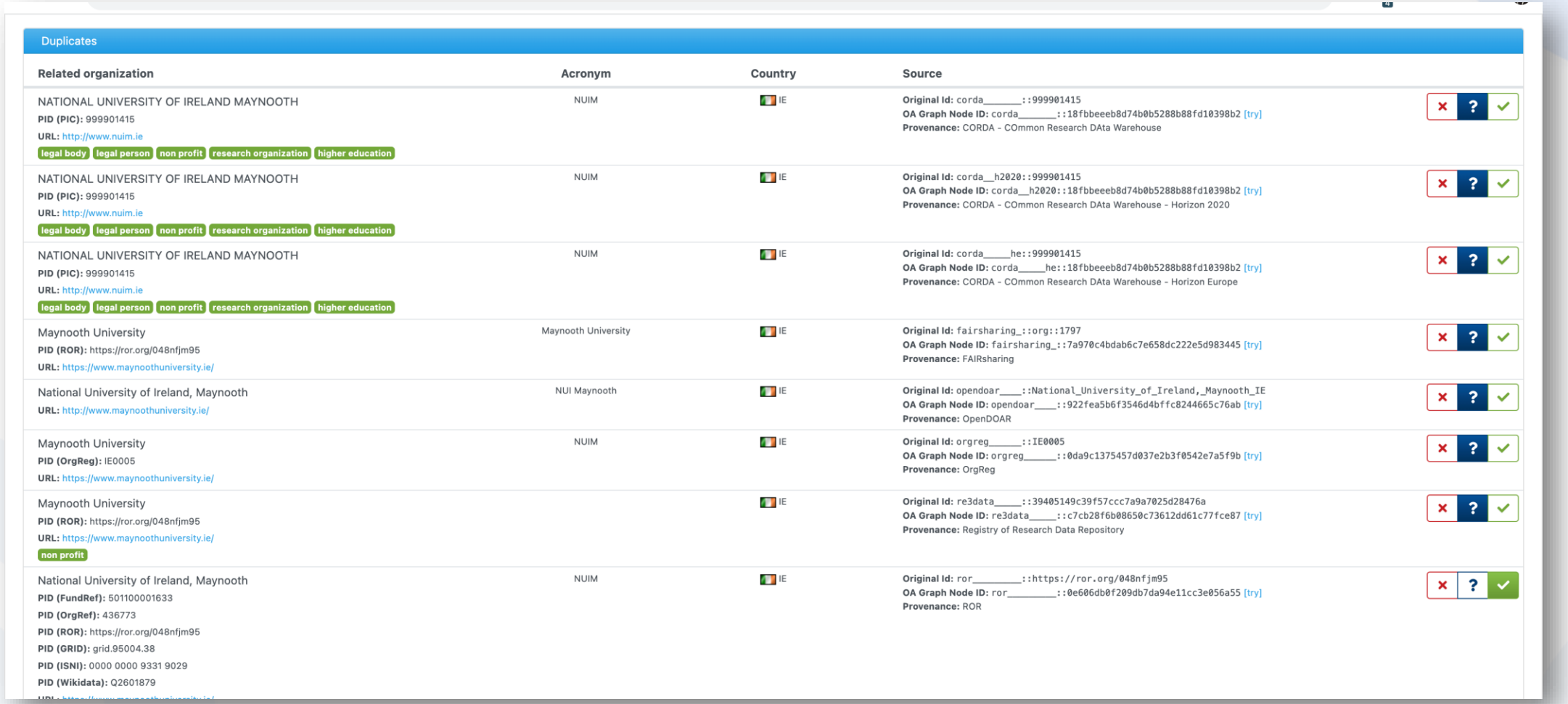
- **Log in** (EDUGAIN, ORCID, Google, GitHub, LinkedIn, OpenAIRE) to access the **LINKING** functionality
  - Enables users to associate **research outputs** with **projects**, **communities**, or **other outputs**, enhancing data representation in OpenAIRE Graph and the Monitor
  - Links show up in next OpenAIRE Graph update
  - A **log** will be **publicly visible** with all added links, with **permission** asked during linking final step

LINKING FUNCTIONALITY





OPENORGS: ORGANISATION DISAMBIGUATION



Duplicates

Related organization	Acronym	Country	Source
NATIONAL UNIVERSITY OF IRELAND MAYNOOTH **PID (PIC):** 999901415 **URL:** http://www.nuim.ie legal body · legal person · non profit · research organization · higher education	NUIM	IE	**Original Id:** corda_____::999901415 **OA Graph Node ID:** corda_____::18fbbeeeb8d74b0b5288b88fd10398b2 [try] **Provenance:** CORDA - COmmon Research DAta Warehouse
NATIONAL UNIVERSITY OF IRELAND MAYNOOTH **PID (PIC):** 999901415 **URL:** http://www.nuim.ie legal body · legal person · non profit · research organization · higher education	NUIM	IE	**Original Id:** corda__h2020::999901415 **OA Graph Node ID:** corda__h2020::18fbbeeeb8d74b0b5288b88fd10398b2 [try] **Provenance:** CORDA - COmmon Research DAta Warehouse - Horizon 2020
NATIONAL UNIVERSITY OF IRELAND MAYNOOTH **PID (PIC):** 999901415 **URL:** http://www.nuim.ie legal body · legal person · non profit · research organization · higher education	NUIM	IE	**Original Id:** corda_____he::999901415 **OA Graph Node ID:** corda_____he::18fbbeeeb8d74b0b5288b88fd10398b2 [try] **Provenance:** CORDA - COmmon Research DAta Warehouse - Horizon Europe
Maynooth University **PID (ROR):** https://ror.org/048nfjm95 **URL:** https://www.maynoothuniversity.ie/	Maynooth University	IE	**Original Id:** fairsharing_::org::1797 **OA Graph Node ID:** fairsharing_::7a970c4bdab6c7e658dc222e5d983445 [try] **Provenance:** FAIRsharing
National University of Ireland, Maynooth **URL:** http://www.maynoothuniversity.ie/	NUI Maynooth	IE	**Original Id:** opendoar____::National_University_of_Ireland,_Maynooth_IE **OA Graph Node ID:** opendoar____::922fea5b6f3546d4bffc8244665c76ab [try] **Provenance:** OpenDOAR
Maynooth University **PID (OrgReg):** IE0005 **URL:** https://www.maynoothuniversity.ie/	NUIM	IE	**Original Id:** orgreg_____::IE0005 **OA Graph Node ID:** orgreg_____::0da9c1375457d037e2b3f0542e7a5f9b [try] **Provenance:** OrgReg
Maynooth University **PID (ROR):** https://ror.org/048nfjm95 **URL:** https://www.maynoothuniversity.ie/ non profit		IE	**Original Id:** re3data_____::39405149c39f57ccc7a9a7025d28476a **OA Graph Node ID:** re3data_____::c7cb28f6b08650c73612dd61c77fce87 [try] **Provenance:** Registry of Research Data Repository
National University of Ireland, Maynooth **PID (FundRef):** 501100001633 **PID (OrgRef):** 436773 **PID (ROR):** https://ror.org/048nfjm95 **PID (GRID):** grid.95004.38 **PID (ISNI):** 0000 0000 9331 9029 **PID (Wikidata):** Q2601879	NUIM	IE	**Original Id:** ror_____::https://ror.org/048nfjm95 **OA Graph Node ID:** ror_____::0e606db0f209db7da94e11cc3e056a55 [try] **Provenance:** ROR

OPENORGS: ORGANISATION DISAMBIGUATION

OPEN ORGS

About Support Login Demo version

Unambiguously identifying organizations involved in

2 The same organizations usually appear with a variety of names (e.g., full legal name, short or alternative names, acronym), and different metadata information in different data sources.

OPEN ORGS

About Support Login Demo version

Home > Data curation

Data curation

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est.

Become a curator

OPEN ORGS

Bridging registries of research organizations

Say goodbye to ambiguous research organization data with O
Our powerful tool streamlines information from multiple resea
organization registries, such as ROR and EC, to populate the O
Research Graph with clear and accurate data.

Demo version Become a curator

2

OpenAIRE Guidelines
OpenAIRE PROVIDE

OpenAIRE Interoperability Guidelines

https://guidelines.openaire.eu/

- A common metadata framework for exchanging interoperable metadata information.
- Based on well-documented existing metadata schemas (Dublin Core, Datacite, cerif) and vocabularies (COAR)
- Established to:
  - Support the Open Access strategy of the European Commission
  - Ensure a full alignment with the EOSC Interoperability Framework
  - Meet the requirements of the OpenAIRE infrastructure
- For data sources:
  - Provide orientation to define and implement their local data management policies according to the requirements of OpenAIRE.
  - Assist to expose publications, datasets and CRIS metadata via the OAI-PMH protocol to integrate with OpenAIRE infrastructure.

OpenAIRE | PROVIDE

https://provide.openaire.eu/

- **One-stop-shop** web service
  - *for*
- **Data sources** (institutional & thematic repositories, research data repositories, journal, aggregators, CRIS systems)
  - *who*
- **Interact with OpenAIRE** (share and exchange metadata and content)

- Front-end access to many of OpenAIRE's backend services

SATISFYING THE USER NEEDS

Validate

The **Validator** service assess the metadata quality by measuring your data source compatibility with the **OpenAIRE guidelines** to expose your content by using global standards.

https://guidelines.openaire.eu/

Register

If validation succeeds your data source can be **registered for regular aggregation** and indexing in OpenAIRE. Additionally, you become part of the EOSC ecosystem and a global interlinked network. Reach a wider audience around the world.

Enrich

Improve your metadata. Get more connections. **OpenAIRE Broker** service offers a wealth of information on scholarly communication data. Find out what interests you and subscribe to receive metadata enrichments for your records.

ttps://www.openaire.eu/content-enrichment-guide

Measure

Share usage data via OpenAIRE's global **Metrics Service – usage counts**. Get the benefit of an aggregated environment.

https://usagecounts.openaire.eu

Collection monitor

Collection monitor provides the **aggregation history**. Track the aggregation stage, collection mode and the number of records aggregated thus improving your metadata quality.

Notifications

Once registered in OpenAIRE, data sources can use the PROVIDE Dashboard functionalities and **get notifications & alerts** (enrichments, errors, useful info).

COMPATIBILITY WITH THE OPENAIRE GUIDELINES

- **Improved interoperability** to meet latest IT and repository standards
  - Your content is more:
    - Contextualised (links and relationships to other types of research outcomes and entities),
    - Flexible (using different and improved vocabularies)
    - Embedded in the R&I ecosystem (alignment with Open Science mandates and standards and support for well-established metadata schemas, namespace abbreviations)
- This is your road to **FAIR**ness
  - If you are compliant with the OpenAIRE guidelines

  then
  - you are also FAIR enough
- **EOSC**: OpenAIRE is your entry/gateway to EOSC.
  - If you are compliant with the OpenAIRE guidelines

  then
  - You will be onboarding the EOSC Portal Catalogue and Marketplace integrated platform

COMPATIBILITY WITH THE OPENAIRE GUIDELINES - KEY POINTS TO FOLLOW

- Expose your repository's metadata via an OAI-PMH endpoint in the format addressed by the Application Profile of the OpenAIRE guidelines (for v.4.0 click here)
- Follow the instructions for all the fields
- Populate at least all the mandatory (M) fields and attributes
- More populated fields -> Improved data quality -> Improved results in your RPO Dasboard

- v3.0
  - Title (M)
  - Creator (M)
  - Access Level (M)
  - Publication Date (M)
  - Publication Type (M)
  - Resource Identifier (M)

- v4.0
  - Title (M)
  - Creator (M)
  - Publication Date (M)
  - Resource Type (M)
  - Resource Identifier (M)
  - Access Rights (M)

HOW DO I EXPOSE THE METADATA OF MY REPOSITORY?

- OpenAIRE Guidelines v4.0 publication in Zenodo:
  - https://doi.org/10.5281/zenodo.1299202

  - Documentation
  - Record samples
  - Metadata xsd schemas

OpenAIRE PROVIDE VALIDATION & REGISTRATION (1)

- The Validator service assess the metadata quality by measuring your data source compatibility with the OpenAIRE guidelines
- Run the validator tool to test the compatibility with the OpenAIRE Guidelines by getting a compatibility score
- Detailed report on the number of records passed for each guideline and indicative failed records from your OAI-PMH endpoint with warnings/errors

OpenAIRE PROVIDE VALIDATION & REGISTRATION (2)

- If not registered -> Register your data source in OpenAIRE through OpenAIRE PROVIDE
- At your repository dashboard in OpenAIRE PROVIDE add a new interface with the new “Desired compatibility level”
- Validation & registration process:
  - https://www.openaire.eu/validator-registration-guide
  - https://www.openaire.eu/openaire-guidelines-for-literature-institutional-and-thematic-repositories

CONTENT AGGREGATION

- The aggregation team will evaluate the compatibility with the OpenAIRE Guidelines.
- OAI-PMH Interface of your data source at PROVIDE Dashboard:
  - Desired compatibility level
    - You select the Guidelines version with which your data source is compatible
  - Current compatibility level
    - Compatibility level returned by the Validator tool after your request to register the interface
  - Compatibility override
    - Actual compatibility level according to the validation process by the Aggregation Team

OpenAIRE | PROVIDE

Benefits

- Visible & richer content
- Linked science (Links from/to research products)
- Up-to-date collection (alerts, enrichments, additions)
- Fully embedded in EOSC infrastructure
- Accurate and qualitative metrics

Next Steps

NEXT STEPS

- National Open Access Monitor: Dashboard Manager Application Form
- Release of **pilot** MONITOR on January 10th *can already preview data at explore.openaire.eu*
- Use linking functionality
- Sign up for OpenORGS deduplication sprints (RPO managers only)
  - January, April, June
- Institutional Repositories
  - Follow OpenAIRE Guidelines
  - Register with OpenAIRE PROVIDE

Thank you!

leonidas.pispiringas@openaire.eu