

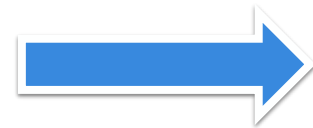
# A user journey in OpenAIRE services through the lens of repository managers

Pedro Príncipe, University of Minho, Alessia Bardi, CNR-ISTI, André Vieira, University of Minho, Jochen Schirrwagen, Bielefeld University,

# WORKSHOP MAIN TOPICS – OpenAIRE SERVICES

---

Interoperability guidelines  
OpenAIRE Research Graph  
Content acquisition policy



---

Literature guidelines V4  
Explore service (demo)  
Infrastructure novelties

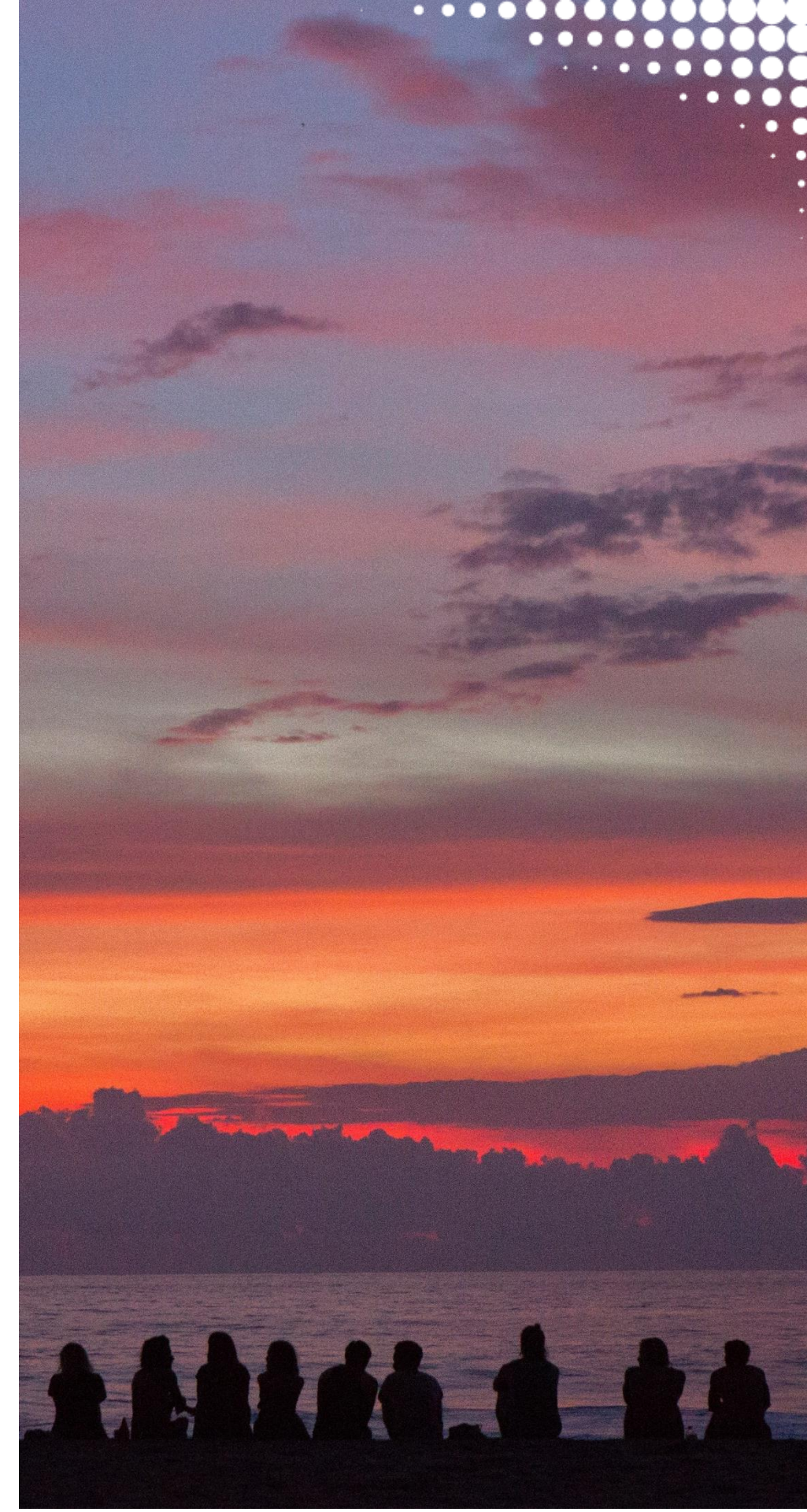
---

Services for Repo Managers  
Dashboard for Content  
Providers  
Broker & usage statistics



---

**Dashboard demo**  
**Service test drive**  
**User feedback**



# AGENDA

**09:00 – Welcome and introduction, Pedro Príncipe**

**09:20 – OpenAIRE graph expansion: an academic graph aggregating all information required to deliver monitoring tools**

**09:50 – OpenAIRE content acquisition policy and the new terms of agreement for content providers.**

**10:05 – Explore service demo (and beta test drive) + Showcase metadata quality issues**

**10:30-11:00 – Coffee break**

**11:00 – OpenAIRE interoperability guidelines overview**

**11:10 – Guidelines for Literature Repositories: implementation and early adopters**

**11:30 – RCAAP use case & HAPLO use case**

**11:50 – OpenAIRE Validator demo - testing the compliance against version 4**

**12:00 – Breakout groups - discussion**

**12:20 – Wrap-up**

**SLIDES HERE:**

**[bit.ly/openaire\\_or2019](https://bit.ly/openaire_or2019)**

<http://box.openaire.eu/index.php/s/V9xSCykE5okIxMIC>

# Workshops Topics – 1st part

---

- 1) OpenAIRE graph expansion: an academic graph aggregating all information required to deliver monitoring tools
- 2) OpenAIRE content acquisition policy and the new terms of agreement for content providers.

>>> **Explore service demo (and beta test drive)**

>>> **Showcase metadata quality issues**

---

- 3) OpenAIRE interoperability guidelines overview
- 4) Guidelines for Literature Repositories: implementation and early adopters: RCAAP use case & HAPLO use case

>>> **Validator demo - testing the compliance against guidelines V4**

>>> **Breakout groups**



# Building the OpenAIRE research graph and the Dashboard services

**CONNECT**

Research communities

**EXPLORE**

Researchers (All)

**PROVIDE**

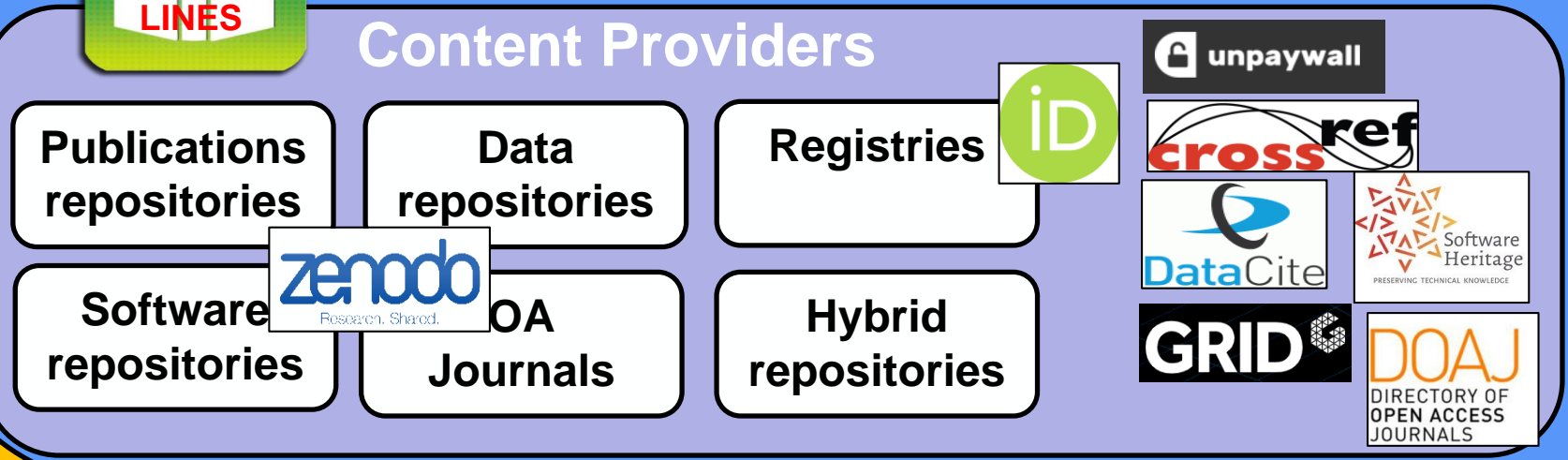
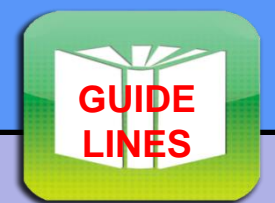
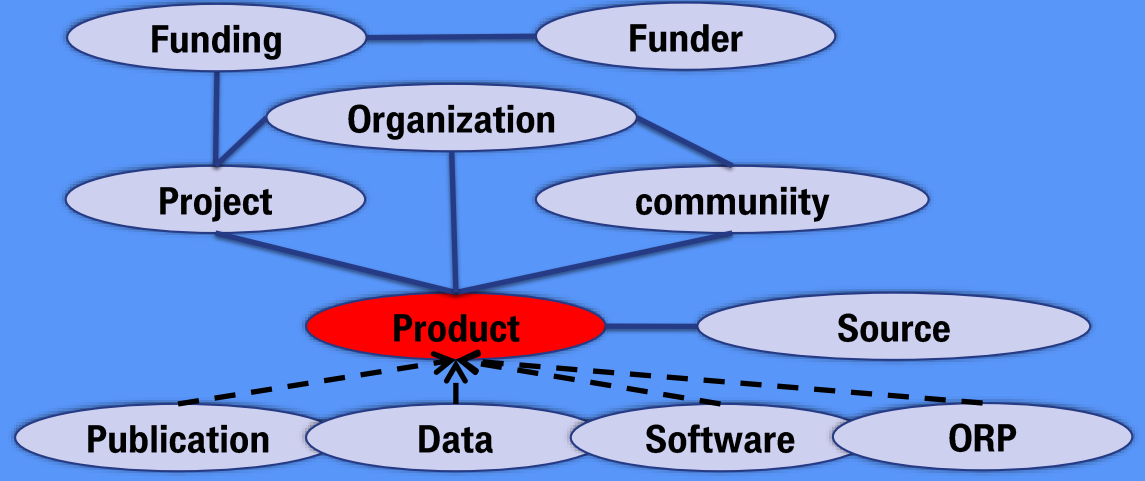
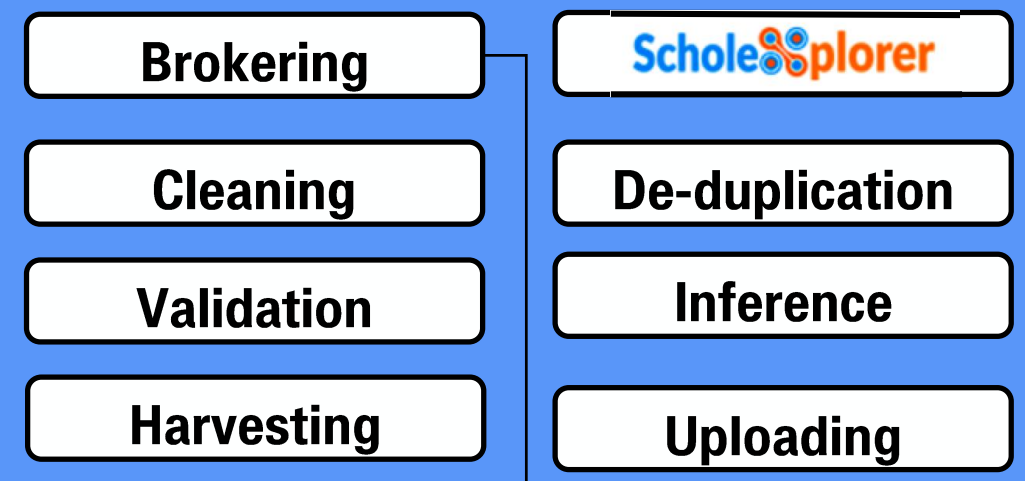
Content providers

**DEVELOP**

Innovators

**MONITOR**

Research managers  
Funders



**Infrastructure**

# OpenAIRE Services

From basic infrastructure level to value added



Search & Share



OpenAIRE | MONITOR



Helpdesk



Training



Research Analytics



Institution Monitoring Dashboard



OpenAIRE | CONNECT



OA Broker



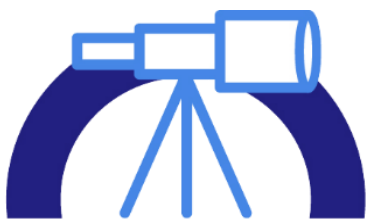
Usage Analytics



OpenAIRE | EXPLORE



Funder Monitoring Dashboard



Open Science Observatory



OpenAIRE | PROVIDE



APIs



Interoperability Guidelines



OpenAIRE | DEVELOP



Zenodo



Validator



easyDMP



Anonymization

OpenAIRE | EXPLORE

SEARCH SHARE LINK CONTENT PROVIDERS SIGN IN

Search in OpenAIRE for scholarly works

26 mi publications, 1 mi research data  
from 15 K content providers and 18 funders  
linked together for an integrated research  
search

<https://explore.openaire.eu>

<https://beta.explore.openaire.eu>

OpenAIRE | EXPLORE

SEARCH SHARE LINK CONTENT PROVIDERS SIGN IN

Search in OpenAIRE for scholarly works

94 mi publications, 8 mi research data  
from 9,800 content providers and 28 funders  
linked together for an integrated research  
search





[www.menti.com](http://www.menti.com)

192517



1

# OpenAIRE graph expansion

*an academic graph aggregating all information required to deliver monitoring tools*

# OpenAIRE

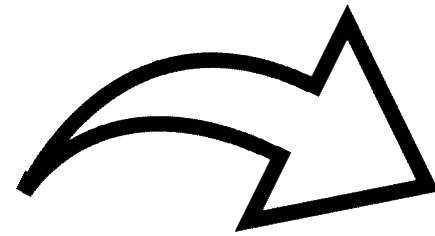
# graph expansion

An academic graph aggregating all information required to deliver monitoring tools

- Slides by Paolo Manghi  
Presented by Alessia Bardi

- *Institute of Information Science and Technologies - CNR*

# Science publishing, evaluation, and monitoring



# Open Science



# Scopus®

The cover of the "Open Science Monitor" report. It features the European Commission logo at the top left, the title "OPEN SCIENCE MONITOR" in blue, and the subtitle "UPDATED METHODOLOGICAL NOTE" in blue. Below the subtitle is the Elsevier logo and the date "Brussels, October 8th 2018". The consortium partners are listed as "The Lisbon Council", "ESADE Business School", and "Centre for Science and Technology Studies (CWTS) at Leiden University". The subcontractor is listed as "Elsevier".

Commissione europea

## OPEN SCIENCE MONITOR

---

### UPDATED METHODOLOGICAL NOTE

Brussels, October 8th 2018

**ELSEVIER**

Consortium partners: The Lisbon Council  
ESADE Business School  
Centre for Science and Technology Studies (CWTS)  
at Leiden University

Subcontractor: Elsevier

# The OpenAIRE research graph

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



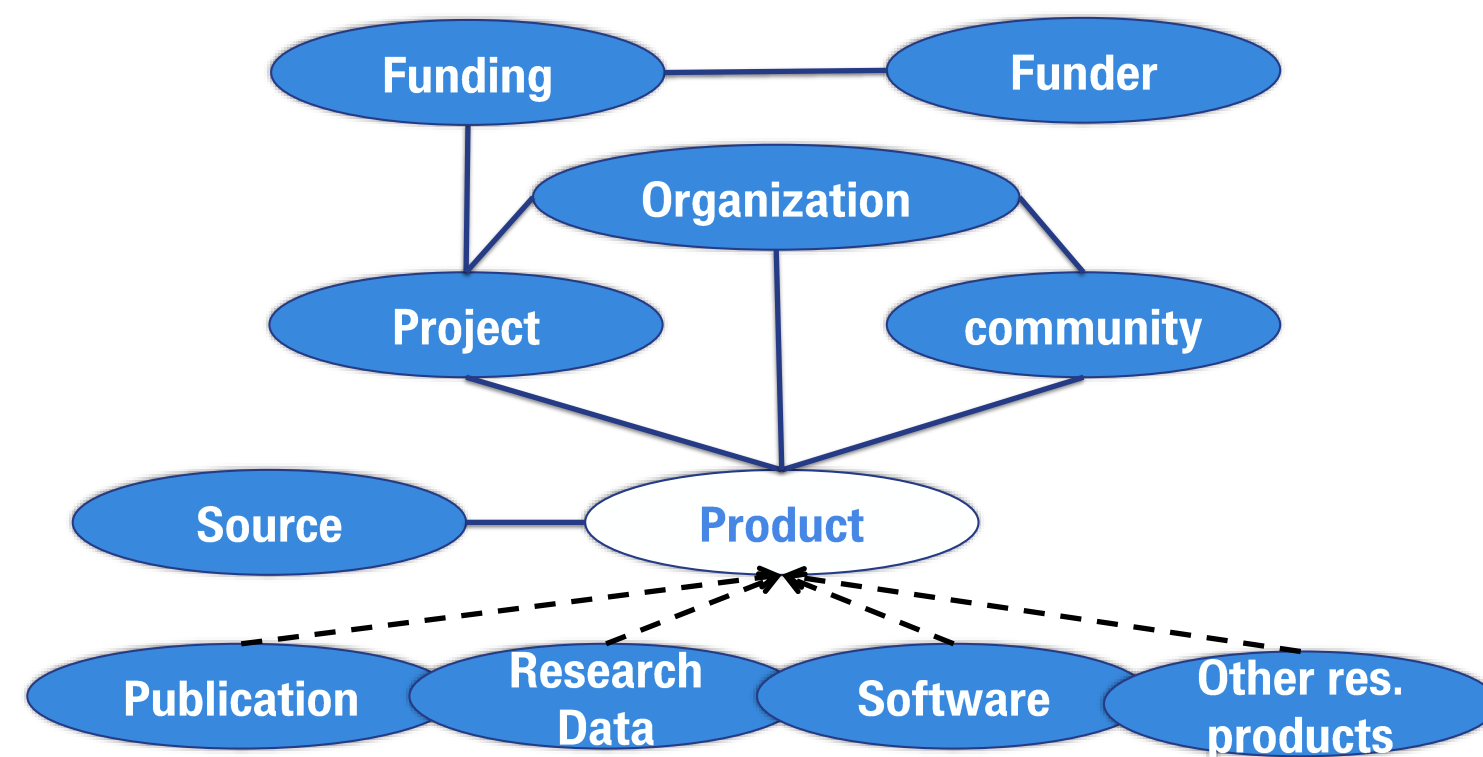
# OpenAIRE e-infrastructure

Materializing the Open Science graph

- Harvested data sources
  - 10K +
- Harvested records
  - 500Mi +
- Publication full-texts
  - 7.5Mi (soon 10.5Mi+)
- Harvested/mined links
  - 200Mi +

**CONNECT** **PROVIDE** **EXPLORE**

**DEVELOP** **MONITOR**



**Deduplication**



**Harvesting**



**Mining**

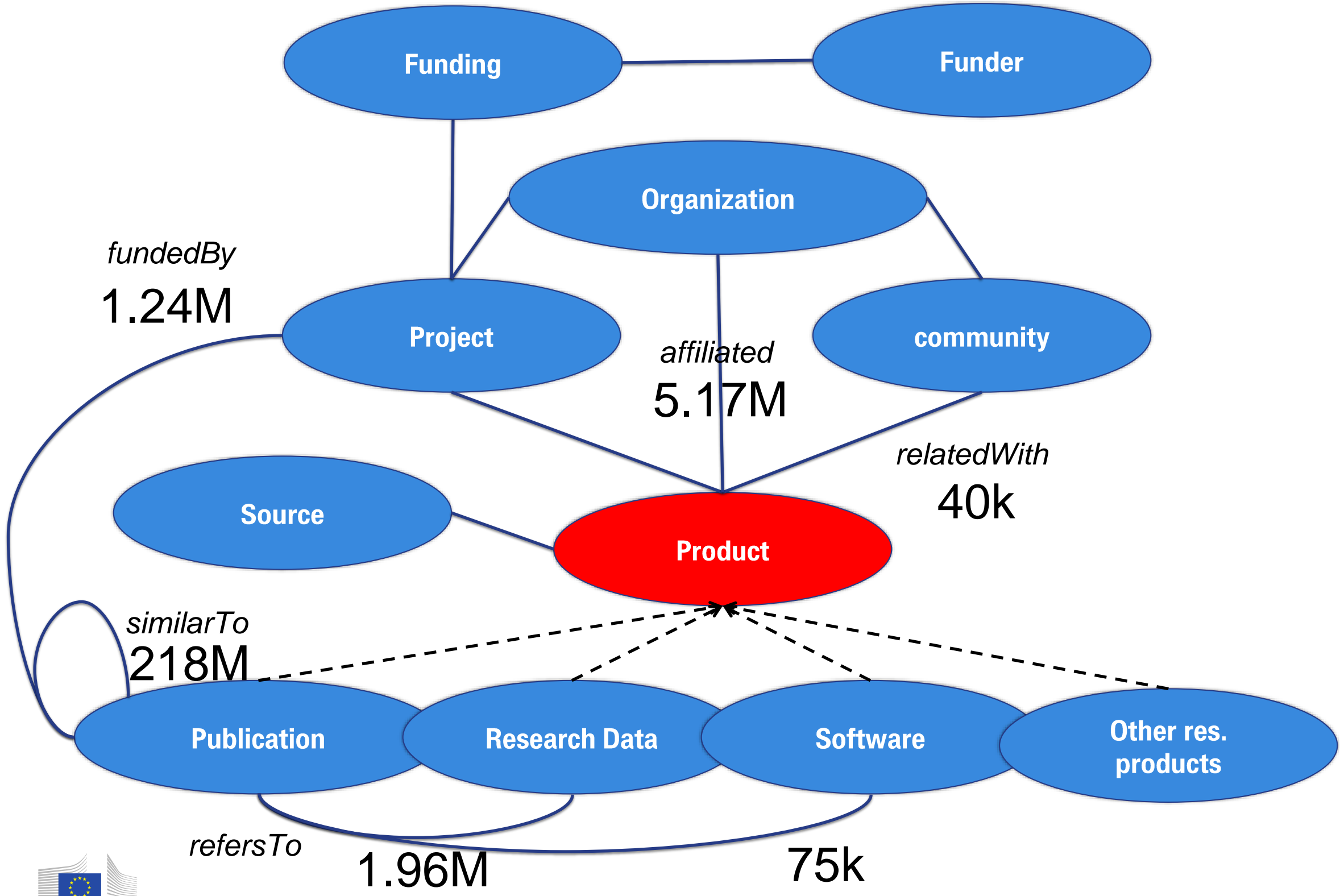


# Open Access articles sources



- **Repositories and publishers**
  - **Download from URLs in harvested metadata: 6.8Mi**
  - **Machine-learning on OA URLs from large aggregators (DOAJ, CrossRef/Unpaywall): 3.3 Mi (downloaded, under integration in BETA)**
- **Publishers metadata/PDFs via CORE-UK**
  - **ResourceSync**
  - **Springer Open Access, etc.: 750K**

# Mining results: links

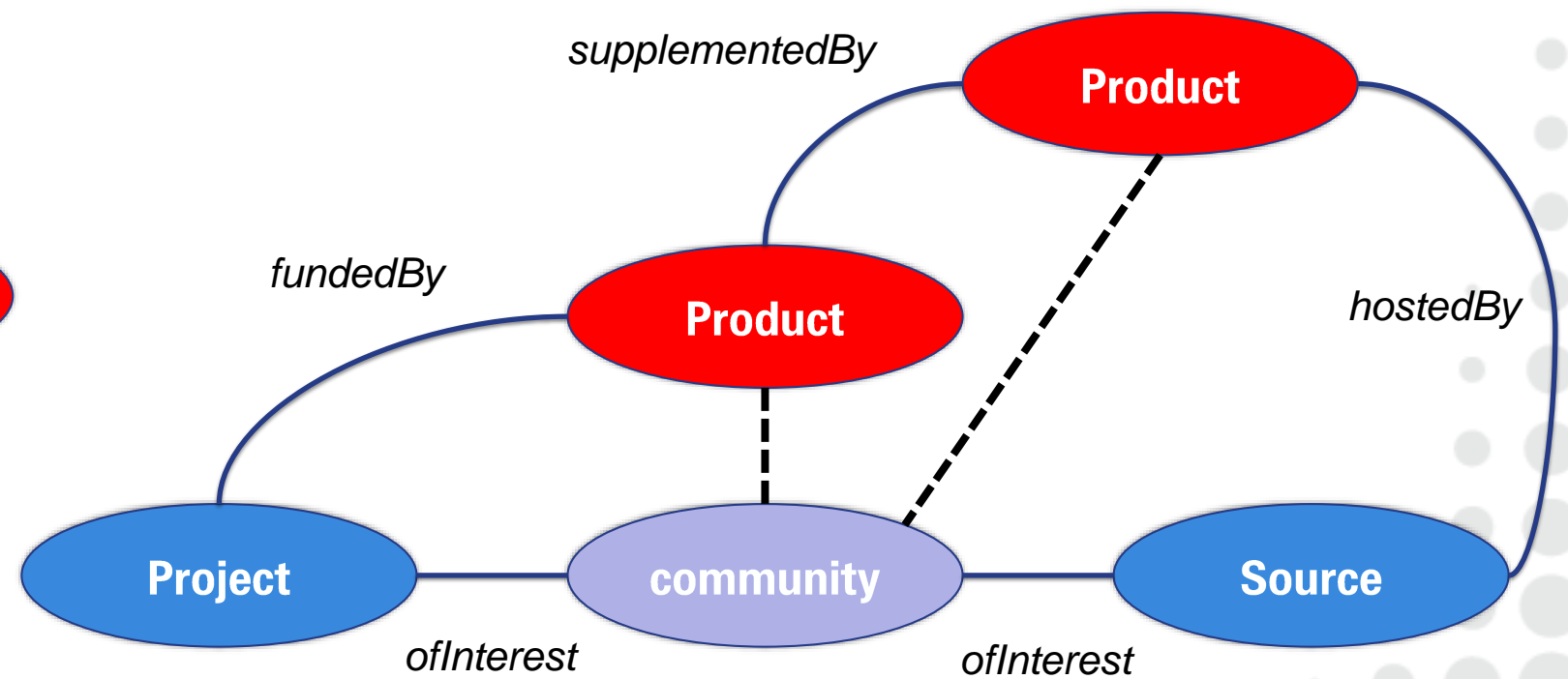
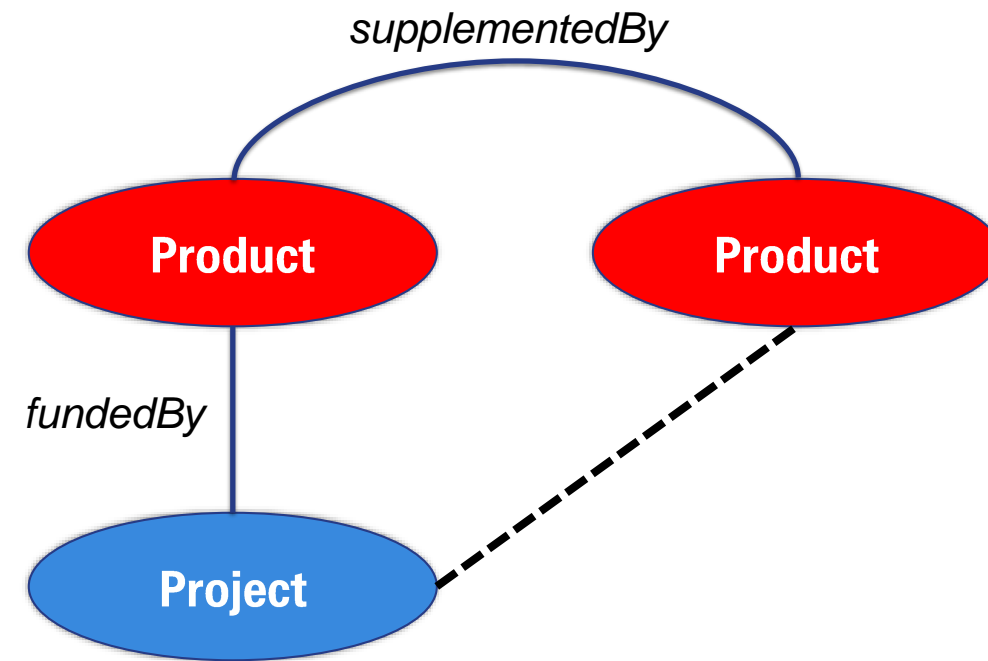
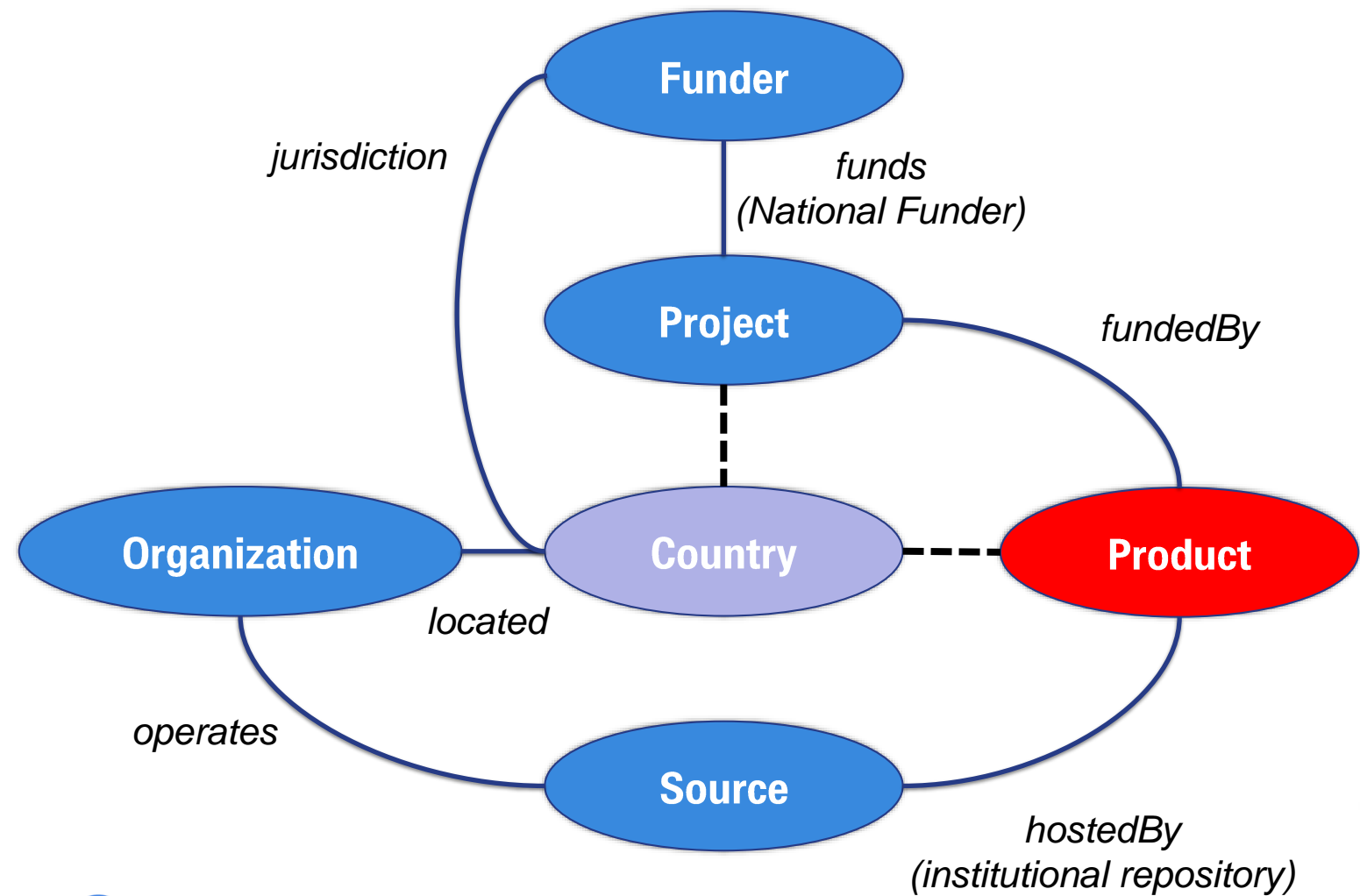




# Mining results: properties

- **Document classification: 3.86M of pubs with at least one class assigned:**
  - **arxiv: 2.35Mi, meshEuroPmc: 3.64Mi, acm: 832k**
- **Document properties**
  - **New abstracts: 1.3Mi**
- **Document references**
  - **168.44M bibliographic references for 5.33M pubs**
- **Document external links**
  - **PDB reference extraction: 320k references (68k of unique pubs)**

# Context Propagation



# Open metadata

Repositories  
Archives



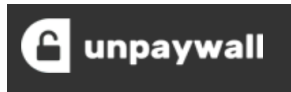
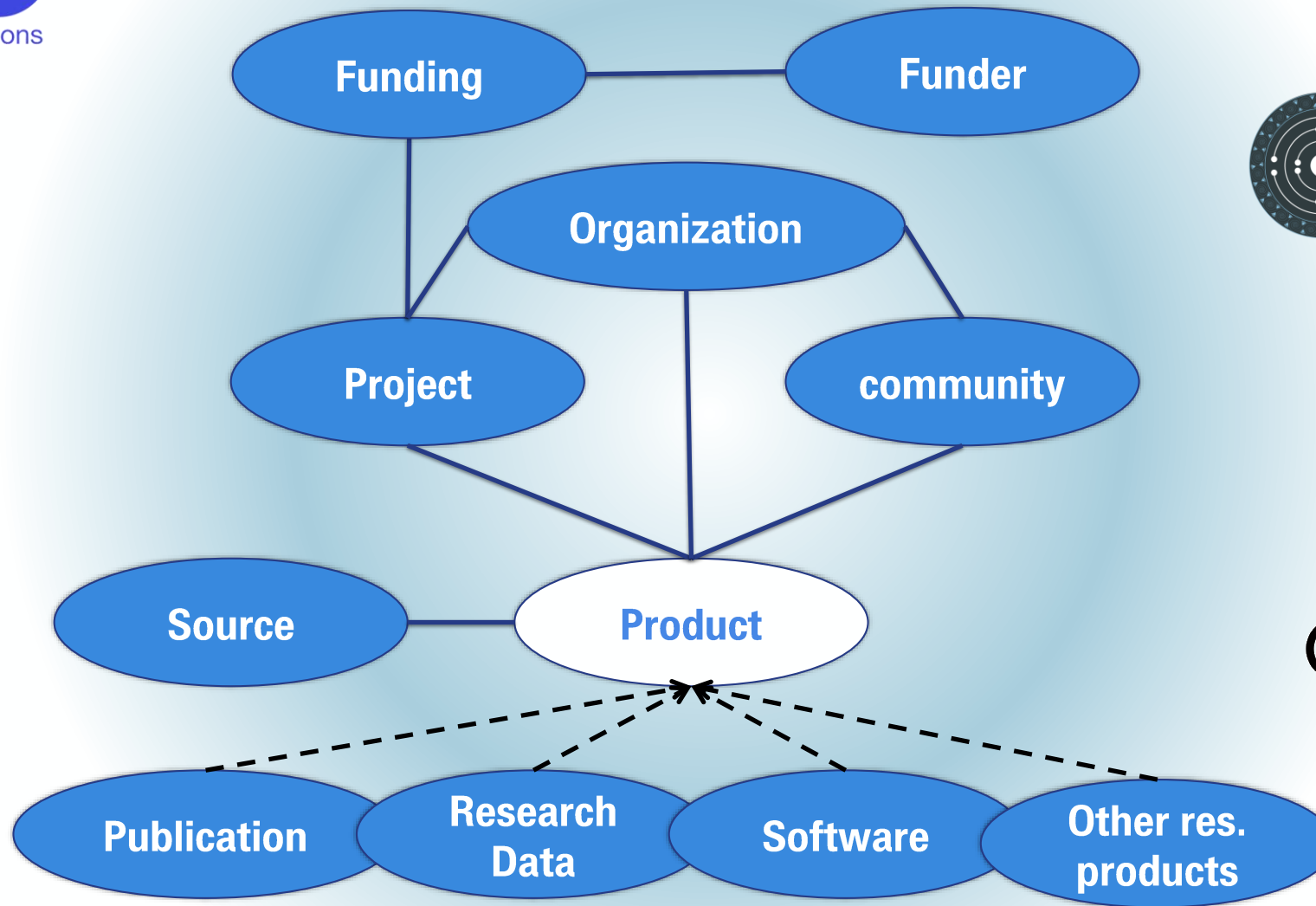
Repositories  
Archives



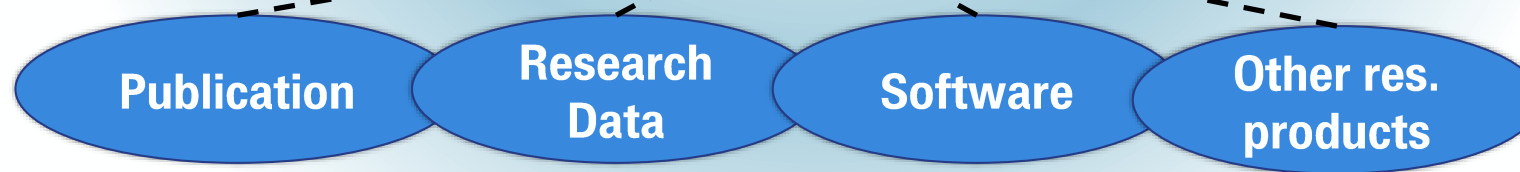
Springer  
Microsoft



# Complete aggregation coverage



... and more



... and more



# De-duplicated

Entity type	# Collected records	# Records after cleaning and de-duplication	# Identified duplicates
Publications	~ 343M	~ 94M	~ 249 millions
Data	~ 5,2M	~4,6M	~ 600K
Software	~150K	~ 134 K	~ 20K
Other	~ 5M	~ 4,5M	~ 500K
Organisations	~ 380K	~220K	~ 160K

More information about the de-duplication framework used by OpenAIRE can be found searching on Zenodo for :

- [“De-duplicating the OpenAIRE Scholarly Communication Big Graph”](#) (poster)
- [“GDup: De-Duplication of Scholarly Communication Big Graphs”](#)



# Participatory

- Rely on quality scholarly communication sources of different kinds
- Include **solutions and content** from any interested and **known** content provider in scholarly communication

Institutional repositories  
Aggregators  
Data archives  
Software repositories  
Research infrastructure sources  
Funder grant databases  
Authors & Orgs entity registries  
Publishers & journals



# Transparent

- Metadata in the graph includes **provenance** when harvested and **reliability indicators** when obtained from mining



# Decentralized

- **Preservation** and **ownership** beyond OpenAIRE
  - Exchanged with other graph initiatives
  - Redistributed via subscription and notification to contributing data sources ([provide.openaire.eu](https://provide.openaire.eu))
- Openly accessible via APIs ([develop.openaire.eu](https://develop.openaire.eu))





# Trusted (in progress)

- Authors in the loop to enrich their ORCID record
- Validation of end-user "claims"

ORCID

Connecting Research  
and Researchers



# Monitoring Tools

# OpenAIRE Open Science Monitoring



**Research  
admins**

**Research impact**

**MONITOR CONNECT**



**Funders**



**Research  
communities**

**Open Access/Science  
trends**

**MONITOR CONNECT**



**Research  
Organization**

# CONNECT

Research Community  
Dashboard for Research  
Infrastructures (Initiatives)

Uptake of Open Science

**MONITOR** publishing practices

Research Impact  
Report to funders

**DEPOSIT ANYTHING**

zenodo

# For building community-specific gateways to Open Science

A virtual environment to implement Open Science publishing practices

All the relevant  
research products

**CONNECT**

... linked

On demand publishing on services of  
Research Infrastructures

Research Community  
Dashboard



# Ongoing collaborations

## Research Infrastructures/initiatives



## Disciplinary research communities

- Sustainable Development Solutions Network (Greece)
- Agricultural and Food Science
- Fisheries and Aquaculture Management
- European Marine Science
- Neuroinformatics
- Digital Humanities and Cultural Heritage

# Aims for Research Infrastructures

## Funding Impact

Publications, research data, software  
published thanks to the existence of the  
RI

## Open Access/Science Impact

Monitoring of Open Science impact:  
data/software FAIRness, reproducibility  
trends

# Added value functionalities

- Open Science indicators

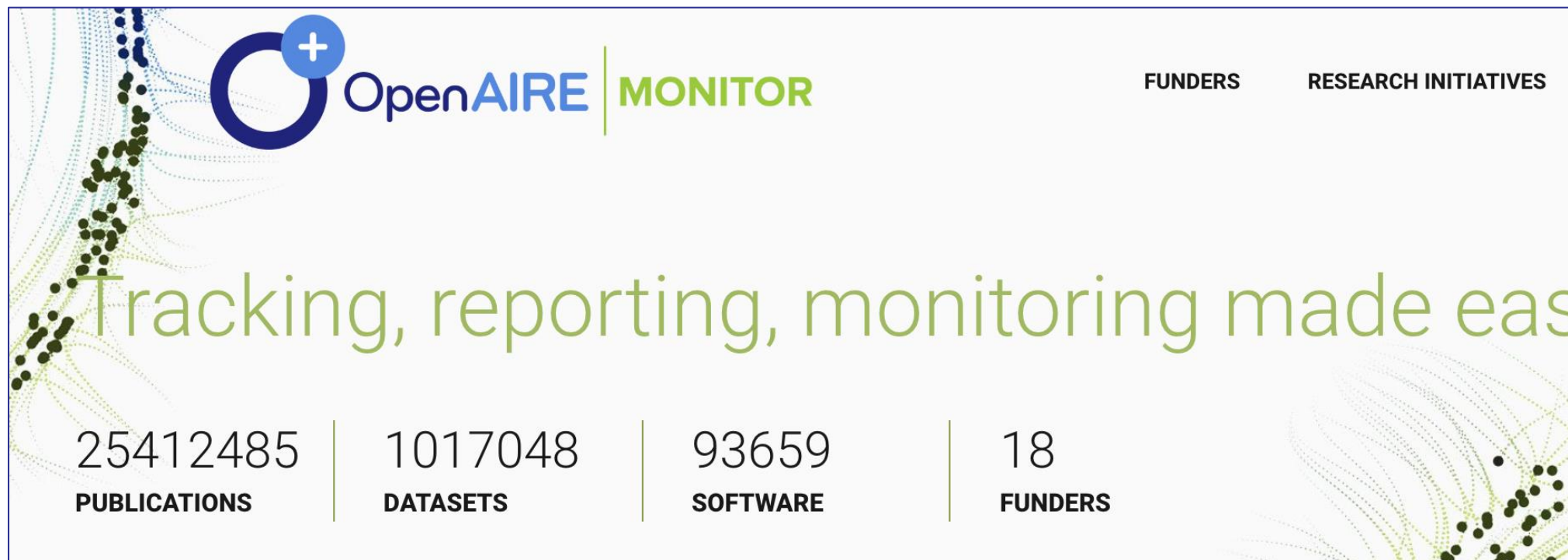




**MONITOR**

Funder Dashboard  
(Project dashboard)

# MONITOR



## European Union

European Commission – H2020  
European Commission – FP7  
ERC - European Research Council (EU)



## International Funders

**Australia**  
ARC - Australian Research Council  
NHMRC - National Health and Medical Research Council Statistics

**USA**  
NIH - National Institutes of Health  
NSF - National Science Foundation



## European Funders

**Austria**  
FWF - Austrian Science Fund

**Croatia**  
HRZZ - Croatian Science Foundation (CSF)  
MZOS - Ministry of Science, Education and Sports of the Republic of Croatia (MSES)

**Finland**  
AKA - Academy of Finland

**France**  
Tara - Tara Expeditions Foundation

**Ireland**  
SFI - Science Foundation Ireland

**Netherlands**  
NWO - Netherlands Organisation for Scientific Research

**Portugal**  
FCT - Fundação para a Ciência e a Tecnologia

**Serbia**  
MESTD - Ministry of Education, Science and Technological Development of Republic of Serbia

**Switzerland**  
SNSF - Swiss National Science Foundation

**Turkey**  
Tubitak - Türkiye Bilimsel ve Teknolojik Araştırma Kurumu

**United Kingdom**  
RCUK - Research Council  
WT - Wellcome Trust

# Aims

## Funding Impact

Publications, research data, software published thanks to grants awarded by the funder

## Open Access/Science Impact

Monitoring of Open Science impact: data/software FAIRness, reproducibility trends

# Added value functionalities

- Funders
  - Trends in research fields: new (multidisciplinary) disciplines
- Projects
  - Interconnections, possible liaisons
- Institutions
  - OA/OS behavior, ability to attract cross-funder grants

**MONITOR**

Institutional Dashboard  
(under development)

# Aims

---

## Research Impact

Ability of researchers affiliated with the institution to produce innovative and quality scientific products

---

---

## Service capacity impact

Ability of services maintained and operated by the institution to support researchers at producing or storing scientific products

---

---

## Funding impact

Ability of institution to reach funding from different funders and disciplines

---

# Added value functionalities

- Funders
  - Recent and past EC and other funders' activities (representing various funding levels)
  - Checking compliance to funder mandates
- Institutions
  - Collaboration network (by institution) via projects and products
- Projects
  - Compare project portfolio against that of other similar institutions (anonymized?)

*Thought about an added value functionality we did not mention?*

Tell us!

<https://beta.explore.openaire.eu>

*Explore the BETA graph and tell us how to improve!*





**Questions?**

A silhouette of five people standing on a cliff edge, celebrating with their arms raised against a sunset sky. The scene is dark, with the sun low on the horizon, creating a gradient from orange to blue.

2

## OpenAIRE content acquisition policy

*new terms of agreement for content providers*

**ALL Literature, Research data,  
Software, Other research products**

**[www.openaire.eu/policies](http://www.openaire.eu/policies)**

**Open Access & non-Open Access material**

# OpenAIRE Content Acquisition Policy + complete

ALL Literature, Research data, Software, Other research products

- Respecting the OpenAIRE guidelines (DataCite metadata)
- Using PIDs with resolvers

The screenshot shows the OpenAIRE website's content acquisition policy page. The header includes the OpenAIRE logo and navigation links for SERVICES, SUPPORT, and OPEN SCIENCE IN EUROPE. The main heading is "Content acquisition policy". Below it, there are social media icons and a section titled "AGGREGATION POLICIES BY TYPE OF PRODUCT" with sub-sections: "RATIONALE", "RESEARCH PRODUCTS AND THEIR ASSOCIATED TYPES IN OPENAIRE", and "AGGREGATION POLICY BY CATEGORY OF REPOSITORY". The main content area is titled "Aggregation policies by type of product" and contains three sections: "Literature, Datasets, Software, other research products", "Accession numbers", and "Full-text of scientific literature".

The screenshot shows the Zenodo page for the OpenAIRE Content Acquisition Policies. The header includes the Zenodo logo, a search bar, and navigation links for Upload and Communities. The main heading is "OpenAIRE Content Acquisition Policies" with a date of October 5, 2018, and a green "Open Access" badge. The page shows 413 views and 371 downloads. The authors listed are Becker, Amelie; Loeden, Aenne; Manghi, Paolo; Principe, Pedro; Schirrwagen, Jochen. The document is indexed in OpenAIRE. The publication date is October 5, 2018, and the DOI is 10.5281/zenodo.1446408. The keywords are content, acquisition, policy, data, scholarly communication, and OpenAIRE. The grants listed are European Commission, OpenAIRE-Advance - OpenAIRE Advancing Open Scholarship (777541), and OpenAIRE-Connect - OpenAIRE - CONNECTing scientific results in support of Open Science (731011). The license is Creative Commons Attribution 4.0 International.

released 05-Oct-2018,

<https://doi.org/10.5281/zenodo.1446408>

**ALL SCIENTIFIC RESEARCH PRODUCTS**

literature, dataset, software,  
other research products

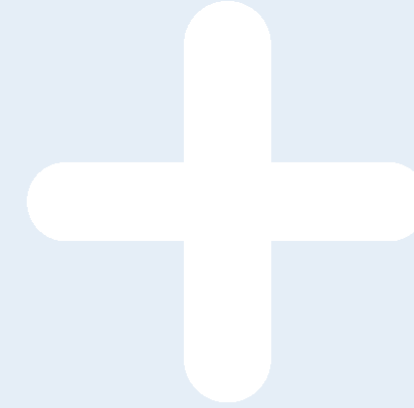
**OF ALL ACCESS LEVELS**

open, closed, metadata only

**METADATA QUALITY**

with a minimal quality  
conditions under which  
metadata can be accepted

**what data/metadata  
we collect**



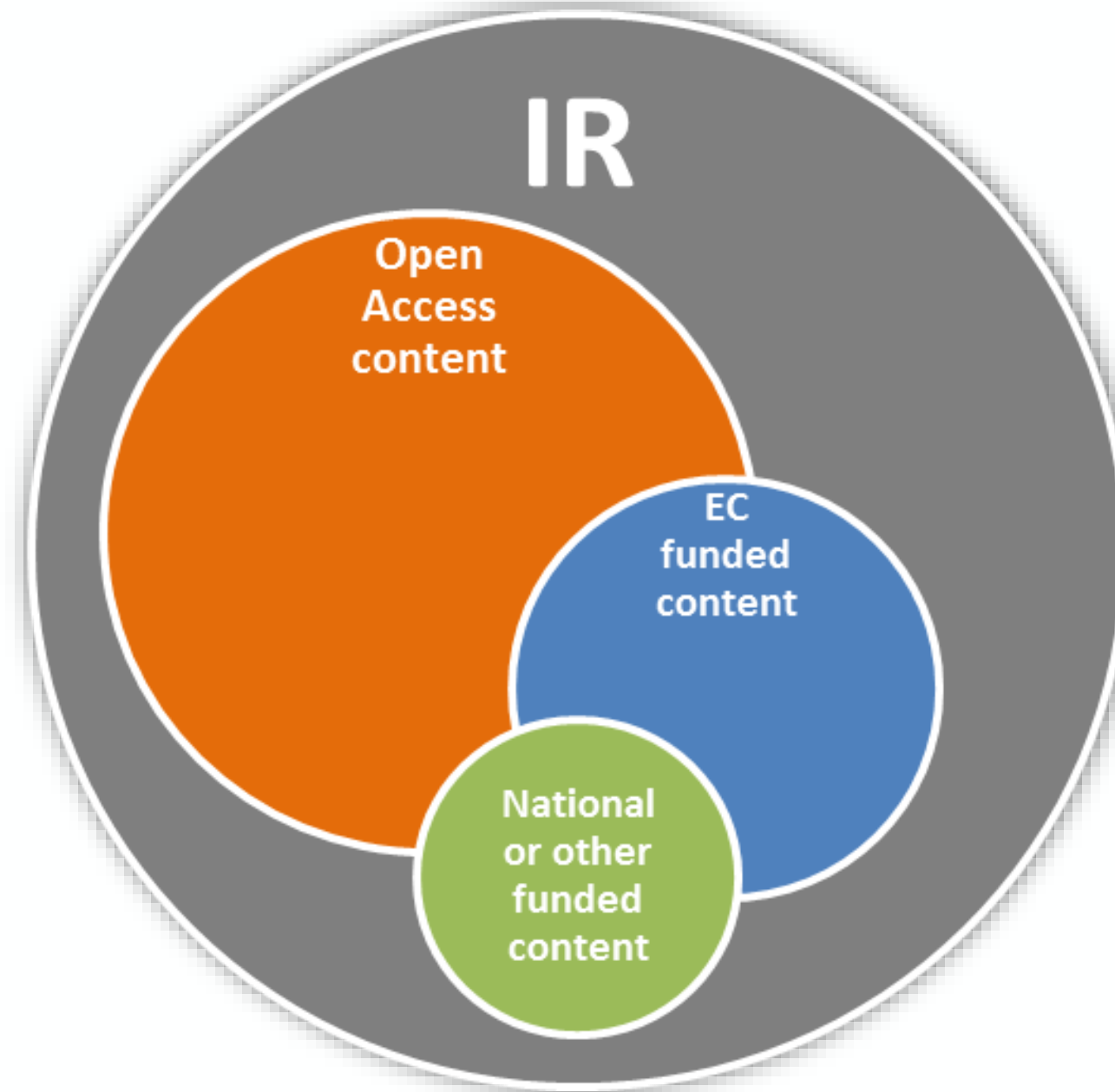


**how we process**

It's important that the access level of a record is made clear

Is vital that the access level of a record is clear (by an access level statement on record level, alternately by the use of specific OAI-sets)

Each record must contain a PID (or URL) that resolves to a splash page



**Metadata describing Open Access and non-Open Access material will be included and links to other products will be resolved where this is possible.**

[www.openaire.eu](http://www.openaire.eu)

“

Metadata describing Open Access and non-Open Access material will be included and links to other products will be resolved where this is possible (i.e. the provided PIDs have a resolver).

*as stated in the Content Acquisition Policy, published Oct. 2018*

<https://doi.org/10.5281/zenodo.1446408>

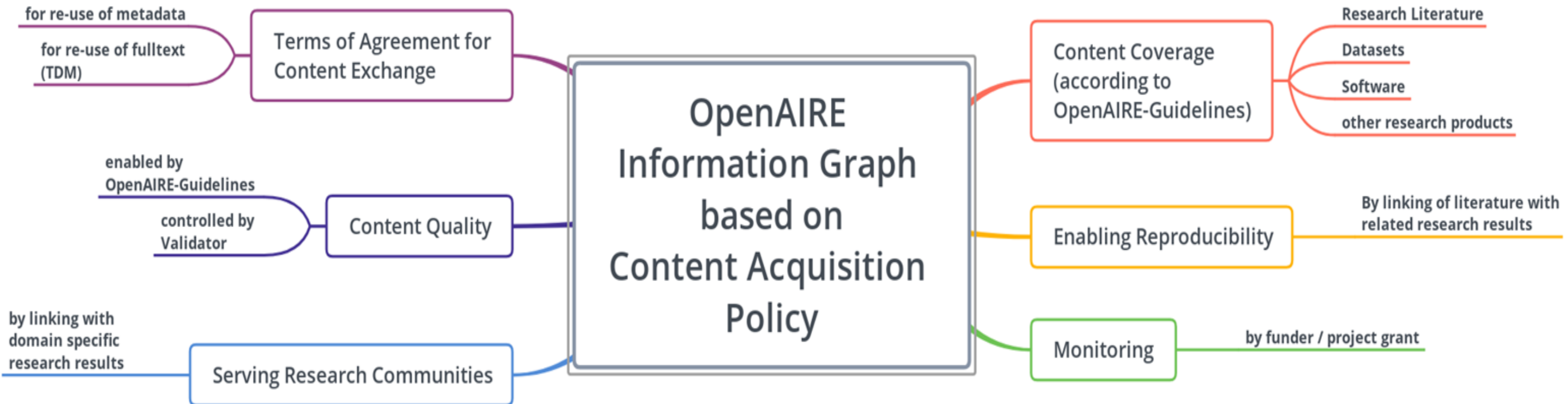
”



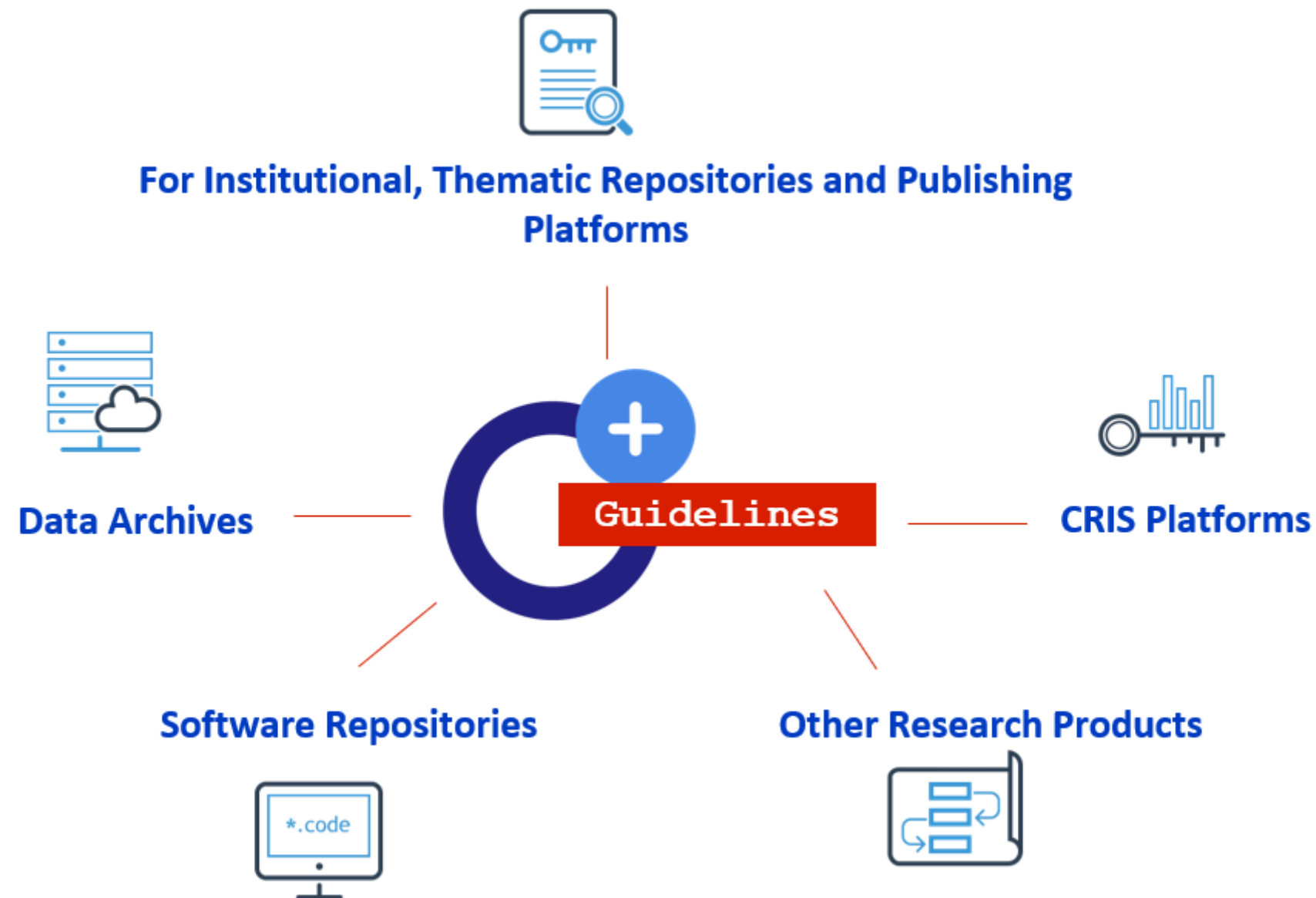
# Objectives of OpenAIRE's Aggregation Policy



# Objectives of OpenAIRE's Aggregation Policy



# CAP supported by the set of Guidelines for Open Science Content Providers



OpenAIRE | EXPLORE

SEARCH SHARE LINK CONTENT PROVIDERS SIGN IN

Search in OpenAIRE for scholarly works

26 mi publications, 1 mi research data  
from 15 K content providers and 18 funders  
linked together for an integrated research  
search

<https://explore.openaire.eu>

<https://beta.explore.openaire.eu>

OpenAIRE | EXPLORE

SEARCH SHARE LINK CONTENT PROVIDERS SIGN IN

Search in OpenAIRE for scholarly works

94 mi publications, 8 mi research data  
from 9,800 content providers and 28 funders  
linked together for an integrated research  
search

The screenshot shows the OpenAIRE Explore website. At the top, there is a navigation bar with links for HOME, EXPLORE, PROVIDE, CONNECT, MONITOR, and DEVELOP. The main header features the OpenAIRE logo and the word 'EXPLORE'. Below the header, there is a search bar with the placeholder text 'Search in OpenAIRE for scholarly works' and a search button. A network graph visualization is visible in the background. The main content area displays the text: '26 mi publications, 1 mi research data from 15 K content providers and 18 funders linked together for an integrated research search'. Below this, there are three main sections: 'Explore' (Discover linked open research), 'Share' (Deposit in a repository of your choice), and 'Link' (Connect all your research).

**DEMO**

<https://explore.openaire.eu>

## How and where is my repository represented in OpenAIRE

- Repository landing page, content, figures, metrics, projects/funders and organizations...
- last index update/collection monitor

OpenAIRE | EXPLORE

EXPLORE PROVIDE CONNECT MONITOR DEVELOP

SEARCH SHARE LINK CONTENT PROVIDERS PEDRO PRÍNCIPE

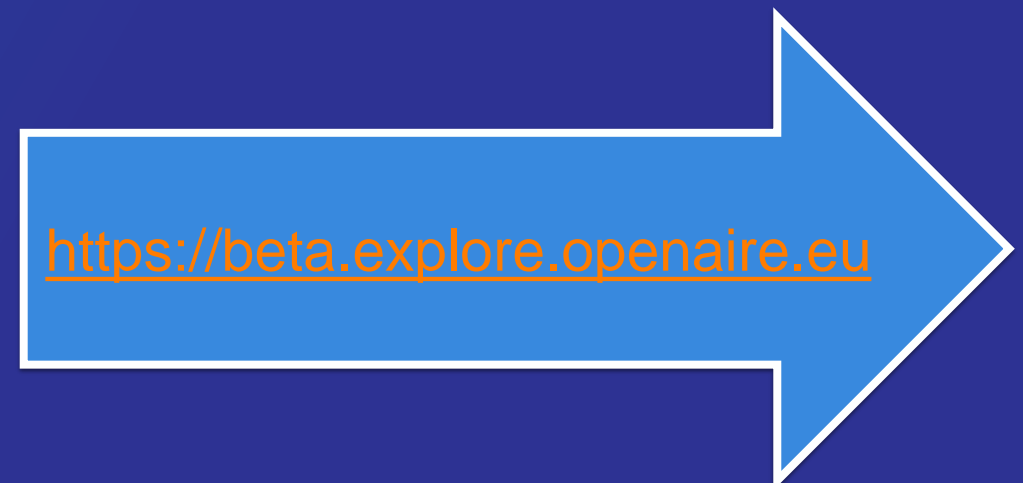
Search in OpenAIRE for scholarly works

94 mi publications, 8 mi research data from 9,800 content providers and 28 funders linked together for an integrated research search

**Explore**  
Discover linked open research  
Search for publications, datasets, software and other research products. See how these are linked together and

**Share**  
Deposit in a repository of your choice  
Select an OpenAIRE compatible repository (2.0+) so that your research is linked to your funding information.

**Link**  
Connect all your research  
If you can't find your research results in OpenAIRE, don't worry!  
Use our **LINK OUT service**, that reaches out to many external



Novelties: orcid, collected from, communities...

# New data sources

- EGI Application database
- OMICS DI
- Kaggle
- ReactToMe
- DOECode
- Unpaywall

# Metadata Quality Challenges

Issue	Affects	Proposed Solutions
Missing values	Indexing, discovery, reuse	Curation by repository team; use OpenAIRE Validator, Broker service
Missing Links and Identifier	Interlinking with other research products; Contextualisation	ScholXplorer, Broker service
Lack of controlled values	Discovery	Use agreed controlled vocabularies according to OpenAIRE Guidelines
Mandatory values only	Discovery and reuse	Broker service



# Records enhanced by de-duplication

- Open Access version coming from one of the sources:

[https://explore.openaire.eu/search/publication?articleId=dedup\\_wf\\_001::0ea9b3d0d7300315854e7f25e499d2b9](https://explore.openaire.eu/search/publication?articleId=dedup_wf_001::0ea9b3d0d7300315854e7f25e499d2b9)

- Document classification:

[https://explore.openaire.eu/search/publication?articleId=od\\_1874::6331f80a2b9758f56609a874e9addc26](https://explore.openaire.eu/search/publication?articleId=od_1874::6331f80a2b9758f56609a874e9addc26)

- For more: look at the Content Provider Dashboard

# Records enhanced by full-text mining (1)

- **Link to software (re-use):**  
<https://explore.openaire.eu/search/publication?articleId=openaire::4405ffb18cc37d73d0daff3650e48f82>
- **Link from a software to its «main» publication:**  
<https://explore.openaire.eu/search/software?softwareId=openaire::949d7264f0efb7a27e521fee9c59209b>
- **Software not available on GitHub, but on SoftwareHeritage only:**  
<https://explore.openaire.eu/search/software?softwareId=openaire::8bf2fbf6cb1f0c9552ca0a6fd0aecfbc>

# Records enhanced by full-text mining

- Reference to a Research infrastructure:  
[https://explore.openaire.eu/search/publication?articleId=nora\\_uio\\_no::3197de1949480eb9f3fc82ba26ad2e25](https://explore.openaire.eu/search/publication?articleId=nora_uio_no::3197de1949480eb9f3fc82ba26ad2e25)
- Link to project:  
[https://explore.openaire.eu/search/publication?articleId=dedup\\_wf\\_001::4be652d611c4bbcf897118bdb564c557](https://explore.openaire.eu/search/publication?articleId=dedup_wf_001::4be652d611c4bbcf897118bdb564c557)

# Quality issues

- Take care of your PIDs:

[https://explore.openaire.eu/search/dataset?datasetId=dedup\\_wf\\_001::69a0263a2925140e015c4470779f79c1](https://explore.openaire.eu/search/dataset?datasetId=dedup_wf_001::69a0263a2925140e015c4470779f79c1)

# A not so rare example

## (Aaltodoc Publication Archive, DSpace)

comparison of OAI-PMH `OpenAIRE endpoint/set` and `standard endpoint/set`

- different number of records (due to former OpenAIRE Content Acqu. Policy)

`completeListSize="12413"` vs. `completeListSize="36886"`

- non-normalised resource types

`<dc:type>info:eu-repo/semantics/article</dc:type>` vs.

`<dc:type>A1 Alkuperäisartikkeli tieteellisessä aikakauslehdessä</dc:type>`

- non-normalised or missing access levels

`<dc:rights>info:eu-repo/semantics/openAccess</dc:rights>` vs.

`<dc:rights>openAccess</dc:rights>`

The background of the slide features a blue-tinted photograph of five people running on a beach. The image is overlaid with a semi-transparent blue filter. The runners are captured in various dynamic poses, conveying a sense of movement and progress. The scene includes a body of water and distant hills in the background.

# 3

## OpenAIRE interoperability guidelines overview

*Guidelines for Literature Repositories: implementation and early adopters*

# OpenAIRE Interoperability Guidelines



**Jochen Schirrwagen**

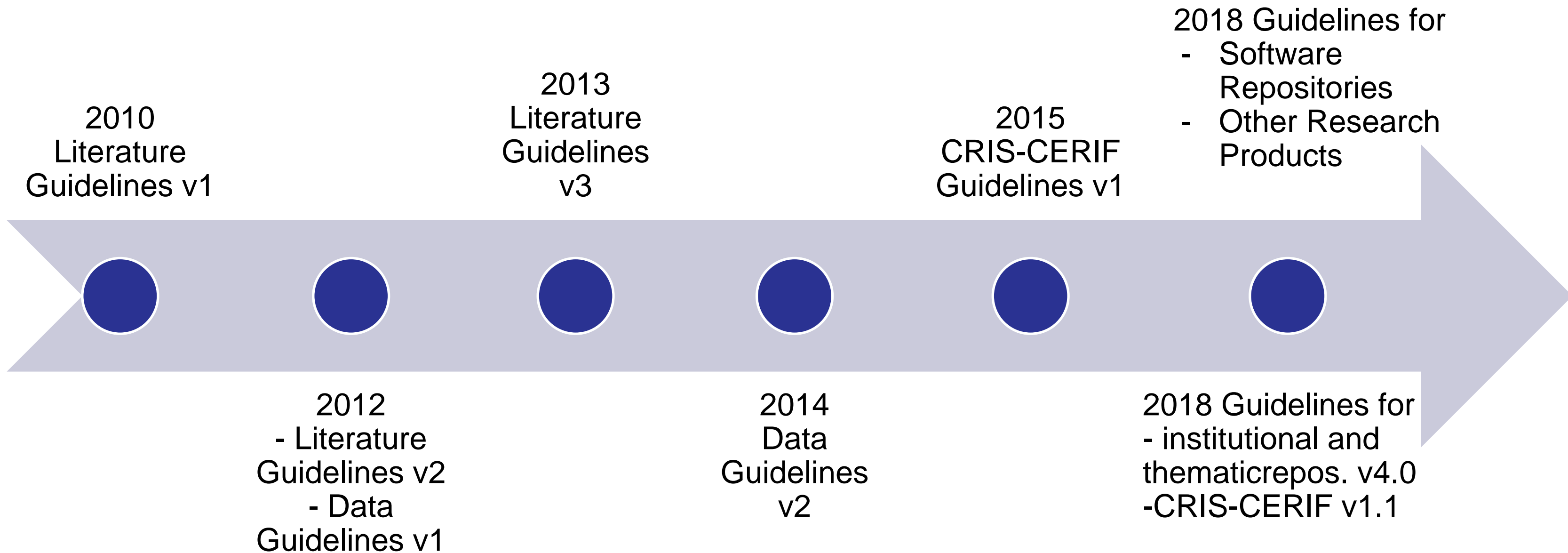
Bielefeld University Library, Germany

Workshop A User Journey in OpenAIRE Services | Hamburg | 10 Jun 2019



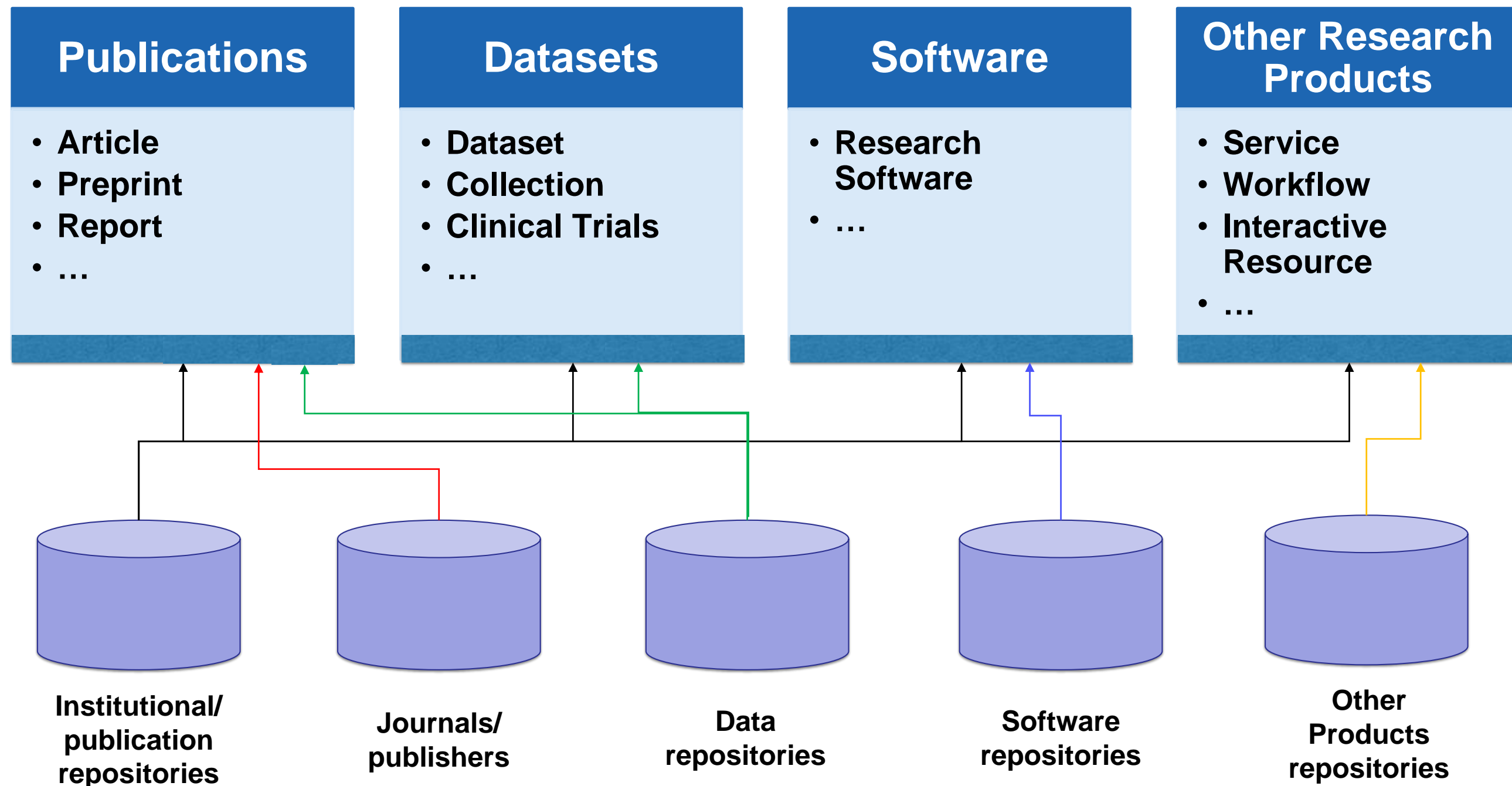
 @openaire\_eu

# Evolution of OpenAIRE-Guidelines





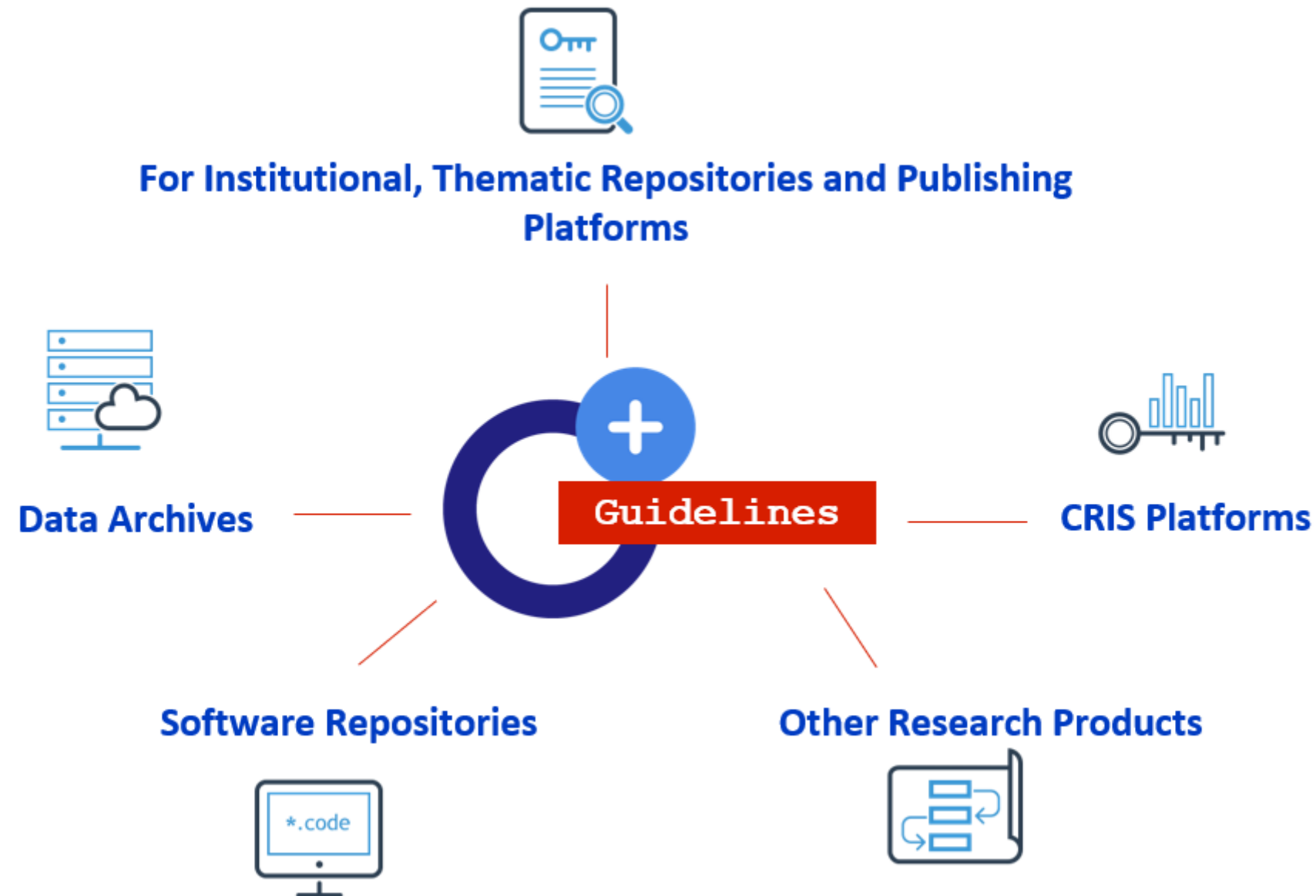
# Diversity of Research Results from Different Types of Sources



# Metadata Goals in OpenAIRE

Goal	Metadata Groups
Discovery and Citability	Descriptive metadata
Accessibility and Reuse	Access Rights, License Conditions
Contextualization	Research Project, Linked Research Artefacts
Interoperability	Identifier for Entities, Controlled Vocabularies
Reporting	Funding Reference
TDM	File Location, License Conditions

# OpenAIRE's Guidelines for Open Science Content Providers



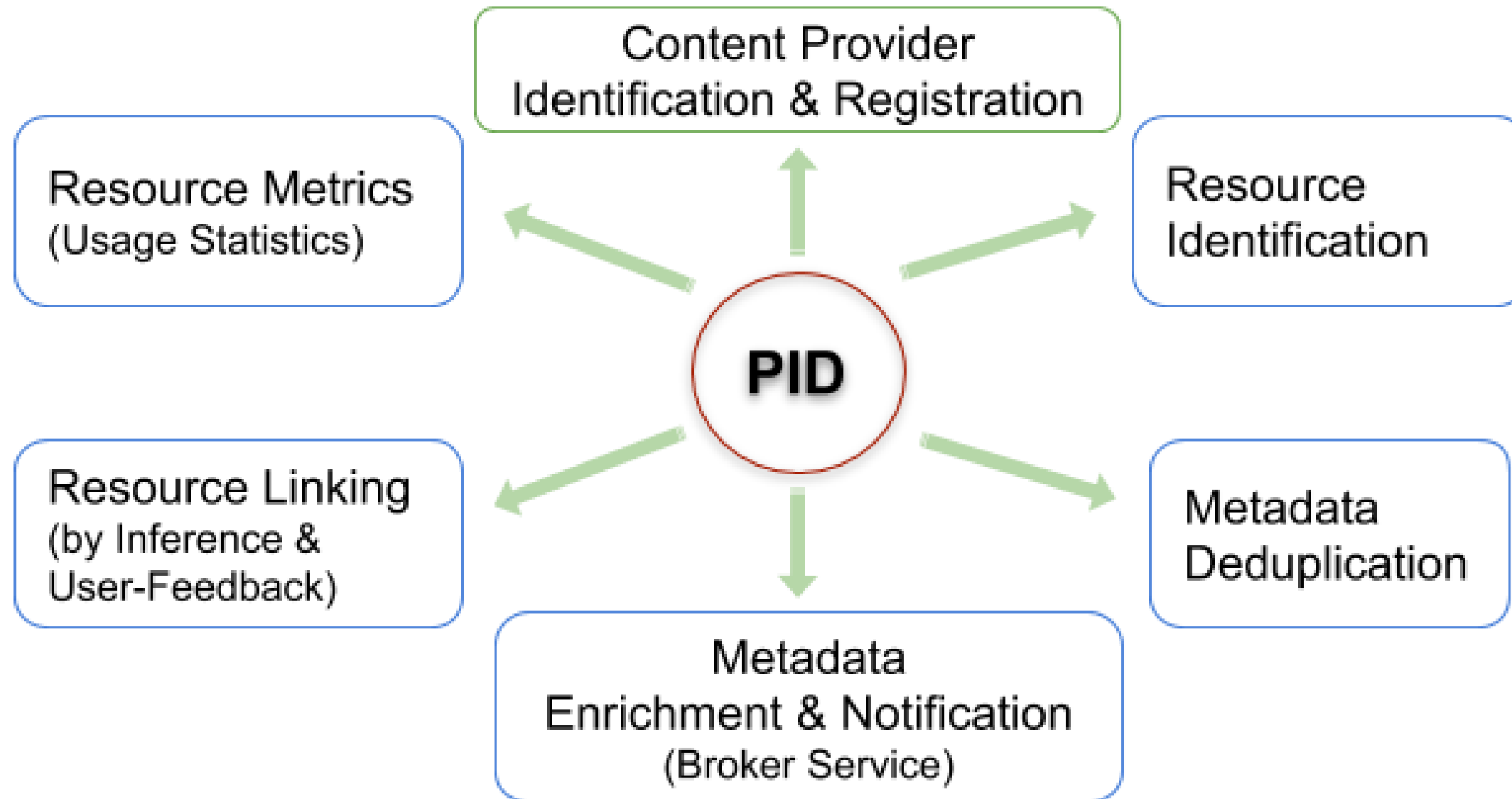
“

**Metadata describing Open Access and non-Open Access material will be included and links to other products will be resolved where this is possible (i.e. the provided PIDs have a resolver).**

*as stated in the  
Content  
Acquisition Policy*

”

# Role of PIDs in OpenAIRE



# OpenAIRE Guidelines for Literature Repository Managers v4.0

- Established standards: Dublin Core and DataCite metadata scheme
- To describe different kinds of scholarly works
- Defines Application Profile
- Controlled Vocabularies and Persistent Identifiers for different entities

<http://dx.doi.org/10.5281/zenodo.1299203>

(Released Nov-2018)

OpenAIRE-Field	Metadata Element	Refinement by Vocabulary
Title (M)	datacite:title	title type
Creator (M)	datacite:creator	name type
Contributor (MA)	datacite:contributor	name type contributor type
Funding Reference (MA)	oire:fundingReference	funderIdentifier type
Alternate Identifier (R)	datacite:alternateIdentifier	alternateIdentifier type
Related Identifier (R)	datacite:relatedIdentifier	relatedIdentifier type relation type resourceType general
Embargo Period Date (MA)	datacite:date	date type
Language (MA)	dc:language	IETF BCP 47, ISO 639-3
Publisher (MA)	dc:publisher	
Publication Date (M)	datacite:date	date type
Resource Type (M)	oire:resourceType	COAR Resource Type Vocabulary
Description (MA)	dc:description	
Format (R)	dc:format	
Resource Identifier (M)	datacite:identifier	identifier type
Access Rights (M)	datacite:rights	COAR Access Right Vocabulary
Source (R)	dc:source	
Subject (MA)	datacite:subject	
License Condition (R)	oire:licenseCondition	
Coverage (R)	dc:coverage	
Size (O)	datacite:size	
Geo Location (O)	datacite:geoLocation	
Resource Version (R)	oire:version	COAR Version Vocabulary
File Location (MA)	oire:file	
Citation Title (R)	oire:citationTitle	
Citation Volume (R)	oire:citationVolume	
Citation Issue (R)	oire:citationIssue	
Citation Start Page (R)	oire:citationStartPage	
Citation End Page (R)	oire:citationEndPage	
Citation Edition (R)	oire:citationEdition	
Citation Conference Place (R)	oire:citationConferencePlace	
Citation Conference Date (R)	oire:citationConferenceDate	
Audience (O)	dcterms:audience	

# Plan S - Requirements and Recommendations

## required

- PIDs for scholarly publications (with versioning)
- Deposition of content with LTP programme (eg. CLOCKSS)
- Article level metadata interoperable non-proprietary format, under a CC0 public domain, incl. funding information
- Machine-readable information on Open Access status and the license



# Plan S - Requirements and Recommendations

## recommended (strongly)

- PIDs for authors (e.g., ORCID), funders, funding programmes and grants, institutions, and other relevant entities.
- Registering self-archiving policy of the venue in SHERPA/RoMEO.
- Availability for download of full text for all publications (including supplementary text and data), eg. JATS XML.
- Direct deposition of publications by the publisher into ... Open Access repositories that fulfil the Plan S criteria.
- OpenAIRE compliance of the metadata.
- Linking to data, code, and other research outputs.
- Openly accessible data on citations according to the standards by the Initiative for Open Citations (I4OC).



# Implementation in Repositories

Software	Supported Version	Status	Comments
DSpace	7 (in prep.) 5 & 6 (in test)	In preparation - DSpace OpenAIRE 4.0 WG Implementations by PT repos RCAAP for v.5	70 days effort (WG timeline plans) Documentation will be available ASAP
EPrints	All	Contacted	May need funding via Jisc or OpenAIRE
Invenio / zenodo		On their roadmap	
Islandora		Contacted	
Librecat		Contacted	
OPUS	4 (in prod.)	Contacted	
MyCoRe		Contacted	
HAL		Contacted	May have very limited resources
Fedora		Will contact	
 		Implemented	

# Implementation Examples



**WELCOME TO ROTHAMSTED REPOSITORY**

This institutional open access repository contains details of the research publications of Rothamsted Research from 1843 to the present day.

Find researchers and their research...

**BROWSE BY RESEARCHER**  
Browse the list of researchers at Rothamsted Research.

**SEARCH THE REPOSITORY**  
Search the repository for outputs matching your chosen query.

**BROWSE BY DEPARTMENT**  
View outputs filtered by Department.

**BROWSE BY YEAR**  
View outputs by year of publication.

This repository supports OAI 2.0 with a base URL of <https://repository.rothamsted.ac.uk/oai2>



```
<OAI-PMH xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/">
  <responseDate> 2019-06-10T05:10:43.510 </responseDate>
  <request verb="GetRecord"> https://repository.rothamsted.ac.uk/oai2 </request>
  <GetRecord>
    <record>
      <header>
        <metadata xmlns:oaire="http://namespace.openaire.eu/schema/oaire/" xsi:schemaLocation="http://namespace.openaire.eu/schema/oaire/ https://www.openarchives.org/OAI/2.0/">
          <oaire:resource xmlns:datacite="http://datacite.org/schema/kernel-4" xmlns:dc="http://purl.org/dc/elements/1.1/" xsi:schemaLocation="http://namespace.openarchives.org/OAI/2.0/">
            <oaire:resourceType resourceTypeGeneral="literature" uri="http://purl.org/coar/resource_type/c_6501"> journal article </oaire:resourceType>
            <oaire:citationTitle> PANS Pest Articles & News Summaries </oaire:citationTitle>
            <oaire:citationVolume> 18 </oaire:citationVolume>
            <oaire:citationIssue> 2 </oaire:citationIssue>
            <oaire:citationStartPage> 186 </oaire:citationStartPage>
            <oaire:citationEndPage> 191 </oaire:citationEndPage>
            <datacite:title>
              The Assessment of Economic Losses in Yield of Annual Crops Caused by Pests, and the Problem of the Economic Threshold
            </datacite:title>
            <datacite:creators>
              <datacite:creator>
                <datacite:creatorName> Judenko, E. </datacite:creatorName>
              </datacite:creator>
            </datacite:creators>
            <datacite:dates>
              <datacite:date dateType="Issued"> 1972-01-01 </datacite:date>
            </datacite:dates>
            <datacite:identifier identifierType="DOI"> 10.1080/09670877209413988 </datacite:identifier>
            <datacite:rights rightsURI="http://purl.org/coar/access_right/c_abf2"> open access </datacite:rights>
          </oaire:resource>
        </metadata>
      </header>
    </record>
  </GetRecord>
</OAI-PMH>
```



# Implementation Examples



University of Liège | Library Network | Login |

Open Repository and Bibliography

Other OA projects at the University Help?

You are here: → ORBI → Browsing ORBI

Search ORBI:

Advanced search

- MyORBI

Browse ORBI by

- Author
- Title
- Issue year
- Journal title
- Document type
- Discipline

ORBI project

- Background
- Content
- Benefits and challenges
- Legal aspects
- Functions and services
- Team

The Open Access movement

- Overview
- Open Access at the University of Liège
- Webbibliography

ORBI is a project of

```
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH" >
  <responseDate>2019-06-10T05:17:07Z </responseDate>
  <request verb="GetRecord" identifier="oai:orbi.ulg.ac.be:2268/179" metadataPrefix="oai_openaire"> http://preprod.orbi.ulg.ac.be/oai/openaire </request>
  <GetRecord>
    <record>
      <header>
        <metadata>
          <oai_openaire xmlns:doc="http://www.lyncode.com/xoai" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:datacite="http://datacite.org/schema/kernel-4" xmlns:oaire="http://namespace.openaire" >
            <oaire:resourceType resourceTypeGeneral="literature" uri="http://purl.org/coar/resource_type/c_6501"> journal article </oaire:resourceType>
            <datacite:rights rightsURI="http://purl.org/coar/access_right/c_abf2"> open access </datacite:rights>
            <oaire:file accessRightsURI="http://purl.org/coar/access_right/c_abf2" objectType="fulltext" mimeType="application/pdf"> https://orbi.uliege.be/bitstream/2268/179/1/Rev%20AIC%2098.pdf </oaire:file>
            <datacite:title xml:lang="fr"> L'ensoleillement et la structure commerciale des places publiques au centre de Liège </datacite:title>
            <datacite:title xml:lang="fr" titleType="TranslatedTitlexsl:text"> Insolation and shops structure on public squares in downtown Liège </datacite:title>
            <datacite:creators>
            <datacite:contributors />
            <datacite:dates>
            <datacite:identifier identifierType="Handle"> info:hdl:2268/179 </datacite:identifier>
            <datacite:alternateIdentifiers />
            <dc:language> fr </dc:language>
            <datacite:relatedIdentifiers>
              <datacite:relatedIdentifier relationType="IsPartOf" relatedIdentifierType="PISSN"> 1140-0307 </datacite:relatedIdentifier>
            </datacite:relatedIdentifiers>
            <datacite:subjects>
              <dcterms:audience xml:lang="en"> Researchers </dcterms:audience>
              <dcterms:audience xml:lang="en"> Behavioral Specialists </dcterms:audience>
              <dcterms:audience xml:lang="en"> Students </dcterms:audience>
            <dc:description xml:lang="en">
              For many years, shop location has been studied at different scales from a socio-economical point of view. Sunshine, on the other hand, is a local phenomenon well-known especially for climatologists and architects. This case-based contribution shows the relationships existing between sunshine and shop location on squares.
            </dc:description>
            <dc:description xml:lang="fr">
              Depuis des années, la localisation des commerces a été étudiée à différentes échelles du point de vue socio-économique. L'ensoleillement quant à lui est un phénomène local bien connu de tout un chacun et en particulier des climatologues et architectes. Cette contribution, basée sur une étude de cas, montre l'existence de relations, sur les places publiques, entre la structure commerciale et l'ensoleillement.
            </dc:description>
            <oaire:citationTitle> Publications de l'Association Internationale de Climatologie </oaire:citationTitle>
            <oaire:citationStartPage> 409-414 </oaire:citationStartPage>
            <oaire:citationVolume> 11 </oaire:citationVolume>
          </oai_openaire>
        </metadata>
      </record>
    </GetRecord>
  </OAI-PMH>
```



# References

- **Guidelines at <https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/>**
- **Schema and examples on github <https://github.com/openaire/guidelines-literature-repositories>**



**Questions?**

# OpenAIRE Guidelines 4 @ RCAAP Project

José Carvalho

[jcarvalho@sdum.uminho.pt](mailto:jcarvalho@sdum.uminho.pt)

# Why?

- Need for a specific format for scientific publications
- More specific fields for metadata fields
- Support for hierarchical information (concept of entities)
- Metadata Aligment with other services (datacite, openaire,....)

# What changes?

- Repositories
- Journals
- Search Portal
- Policies



# What Changes? On Repositories...

- Submission forms

Introduza o nome do(s) autor(es) do trabalho e clique para procurar o respetivo identificador no ORCID ou Ciência Vitae.

**Autores \***



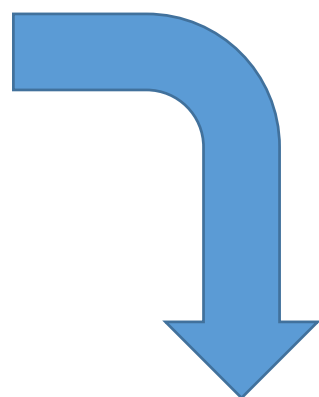
+ Adicionar mais

- Expose in OpenAIRE 4 metadata schema (OAI-PMH)

```
<resource xmlns:vc="http://www.w3.org/2007/XMLSchema-versioning" xmlns:datacite="http://datacite.org/schema/kernel-4"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:doc="http://www.lyncode.com/xoai" xmlns:rdf="http://www.w3.org/TR/rdf-concepts/"
xmlns="http://namespace.openaire.eu/schema/oaire/" xsi:schemaLocation="http://namespace.openaire.eu/schema/oaire/
https://www.openaire.eu/schema/repo-lit/4.0/openaire.xsd">
<datacite:titles>
<datacite:title xml:lang="pt_PT">teste genérico</datacite:title>
<datacite:title xml:lang="pt_PT" titleType="Alternative Title">teste</datacite:title>
</datacite:titles>
<datacite:creators>
<datacite:creator>
<datacite:creatorName>Carvalho, José</datacite:creatorName>
<datacite:givenName>José</datacite:givenName>
<datacite:familyName>Carvalho</datacite:familyName>
<datacite:nameIdentifier nameIdentifierScheme="orcid">0000-0003-1729-3404</datacite:nameIdentifier>
<datacite:nameIdentifier nameIdentifierScheme="ciencialD">6013-8C52-C280</datacite:nameIdentifier>
</datacite:creator>
</datacite:creators>
<file accessRightsURI="http://purl.org/coar/access_right/c_abf2" mimeType="application/pdf"
objectType="fulltext">http://dev4.rcaap.pt/jspui/bitstream/123456789/1597/1/TESTE.pdf</file>
```

# Transformation of project ID into project entity

info:eu-repo/grantAgreement/EC/FP7/612425/EU



```
<fundingReferences>
<fundingReference>
<funderName>European Commission</funderName>
<funderIdentifier funderIdentifierType="Crossref Funder ID">http://doi.org/10.13039/501100008530</funderIdentifier>
<fundingStream>FP7</fundingStream>
<awardNumber>612425</awardNumber>
</fundingReference>
</fundingReferences>
```

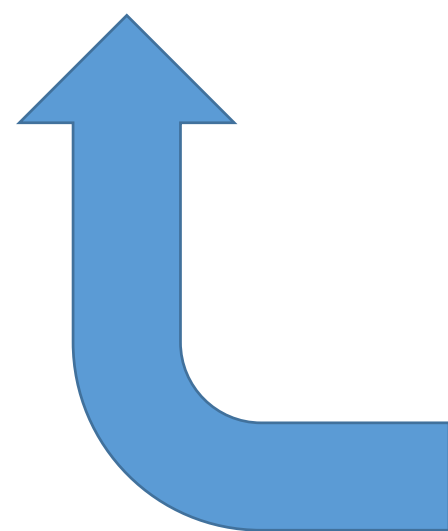
# Submission Forms (authors)

- Author IDs
- COAR Taxonomies

Introduza o nome do(s) autor(es) do trabalho e clique para procurar o respetivo identificador no ORCID ou Ciência Vitae.

**Autores \***

<input type="text" value="Tavares"/>	<input type="text" value="A.C."/>	<input checked="" type="checkbox"/>	<input type="button" value="ID Q"/>
<small>Último nome, ex. Silva</small>	<small>Primeiro Nome, ex. Manuel</small>		<input type="button" value="ID Q"/>



## Pesquisa na autoridade de nomes


Procurar por identificador

**Lista de Resultados**

- Santos, Ana Beatriz
- Santos, Amanda P.
- Santos, Alexandra
- Magalhaes, Carolina
- SANTOS, A. N.
- dos Santos-Ferreira, Adriana M.
- Cancelas, Alberto
- Tavares, A.C.**
- Cristina Corrêa Figueira, Ana
- Santos, Nathália
- Santos, Elivelton
- Santos, Julia
- Santos, Vânia
- Santos, Cibelle

[Apresentar mais resultados](#)



### Informação sobre o autor



Nome  
Tavares, A.C.

Variante  
Tavares, A.C.P.

Afiliação

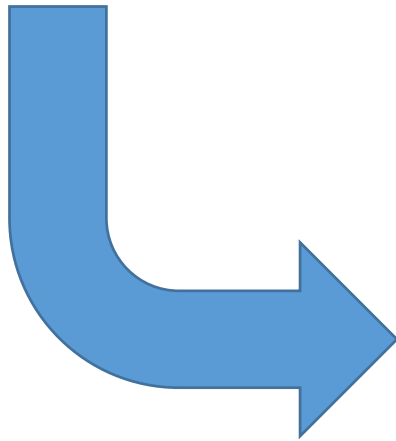
Identificadores  
5E14-3AC4-0733   
0000-0001-8595-218X 

# Expose OAI-PMH with oai\_openaire

oai\_openaire

Namespace <http://namespace.openaire.eu/schema/oaire/>

Schema <https://www.openaire.eu/schema/repo-lit/4.0/openaire.xsd>



```
<resource xmlns:vc="http://www.w3.org/2007/XMLSchema-versioning" xmlns:datacite="http://datacite.org/schema/kernel-4"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:doc="http://www.lyncode.com/xoai" xmlns:rdf="http://www.w3.org/TR/rdf-concepts/"
xmlns="http://namespace.openaire.eu/schema/oaire/" xsi:schemaLocation="http://namespace.openaire.eu/schema/oaire/
https://www.openaire.eu/schema/repo-lit/4.0/openaire.xsd">
  <datacite:titles>
    <datacite:title xml:lang="pt_PT">Embargoed Item Title</datacite:title>
  </datacite:titles>
  <datacite:creators>
    <datacite:creator>
      <datacite:creatorName>Carvalho, José</datacite:creatorName>
      <datacite:givenName>José</datacite:givenName>
      <datacite:familyName>Carvalho</datacite:familyName>
      <datacite:nameIdentifier nameIdentifierScheme="orcid">0000-0003-1729-3404</datacite:nameIdentifier>
      <datacite:nameIdentifier nameIdentifierScheme="ciencialD">6013-8C52-C280</datacite:nameIdentifier>
    </datacite:creator>
  </datacite:creators>
  <datacite:contributors>
    <datacite:contributor contributorType="HostingInstitution">
      <datacite:contributorName>DSpace AuthorProfile</datacite:contributorName>
      <datacite:nameIdentifier nameIdentifierScheme="e-mail"
schemeURI="mailto:webmaster@rcaap.pt">webmaster@rcaap.pt</datacite:nameIdentifier>
    </datacite:contributor>
  </datacite:contributors>
  <fundingReferences>
    <fundingReference>
      <funderName>European Commission</funderName>
      <funderIdentifier funderIdentifierType="Crossref Funder ID">http://doi.org/10.13039/501100008530</funderIdentifier>
      <fundingStream>FP7</fundingStream>
```

# What Changes? On Journals (OJS)...

- OAI PMH (not yet developed)

```
<resource xmlns:vc="http://www.w3.org/2007/XMLSchema-versioning" xmlns:datacite="http://datacite.org/schema/kernel-4"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:doc="http://www.lyncode.com/xoai" xmlns:rdf="http://www.w3.org/TR/rdf-concepts/"
xmlns="http://namespace.openaire.eu/schema/oaire/" xsi:schemaLocation="http://namespace.openaire.eu/schema/oaire/
https://www.openaire.eu/schema/repo-lit/4.0/openaire.xsd">
<datacite:titles>
<datacite:title xml:lang="pt_PT">teste genérico</datacite:title>
<datacite:title xml:lang="pt_PT" titleType="Alternative Title">teste</datacite:title>
</datacite:titles>
<datacite:creators>
<datacite:creator>
<datacite:creatorName>Carvalho, José</datacite:creatorName>
<datacite:givenName>José</datacite:givenName>
<datacite:familyName>Carvalho</datacite:familyName>
<datacite:nameIdentifier nameIdentifierScheme="orcid">0000-0003-1729-3404</datacite:nameIdentifier>
<datacite:nameIdentifier nameIdentifierScheme="ciencialD">6013-8C52-C280</datacite:nameIdentifier>
</datacite:creator>
</datacite:creators>
```

# What Changes? On Search Portal...

- New harvesting process
  - Support for multi metadata schema (oai\_dc; xoai; oai\_openaire)
  - Support for type of resource (repository, journal,...) and type of metadata schema
- New transformation and validation processes
  - DRIVER Types → COAR Types
- New ways to present the information
  - On the user interface
  - On OAI-PMH
  - On REST API

# What Changes? On Policies level...

- New harvesting policy for the National Harvester
- Important for content regulation
- Development of new profile for OpenAIRE 4 on RCAAP Validator
- Important to align with national and international services and developments

# Final Considerations

- Pilot Dspace 5 instance with guidelines OpenAIRE 4 implemented
  - All the information will be available to implemente in Dspace software

## Participation on the Working Group DSpace OpenAIRE

- New harvesting rules and pilot with OpenAIRE 4 soon
- Mappings of information
  - Still some lacks of information between some services (information may be lost)
- Already some suggestions for Guidelines OpenAIRE 4.1



# Thanks!

José Carvalho – [jcarvalho@sdum.uminho.pt](mailto:jcarvalho@sdum.uminho.pt)



[www.menti.com](http://www.menti.com)

192517

# Breakout groups – questions:

- How to optimize the information exchange (metadata and fulltext) between repositories and OpenAIRE? How to reduce the burden to repository managers?
- How could we help - what kind of support to you would like to have from OpenAIRE?
- What are the major metadata quality issues and how to solve them?

# Thank you!

Pedro Príncipe, University of Minho, [pedroprincipe@sdum.uminho.pt](mailto:pedroprincipe@sdum.uminho.pt);

Alessia Bardi, CNR-ISTI, [alessia.bardi@isti.cnr.it](mailto:alessia.bardi@isti.cnr.it)

André Vieira, University of Minho, [andrevieira@sdum.uminho.pt](mailto:andrevieira@sdum.uminho.pt);

Jochen Schirrwagen, Bielefeld University, [jochen.schirrwagen@uni-bielefeld.de](mailto:jochen.schirrwagen@uni-bielefeld.de)  
[info@openaire.eu](mailto:info@openaire.eu)

[info@openaire.eu](mailto:info@openaire.eu)