

Guidelines @ OpenAIRE

for institutional and thematic repository managers

Bielefeld University

FECYT OpenAIRE workshop | León / ES | 25 SEP 2019

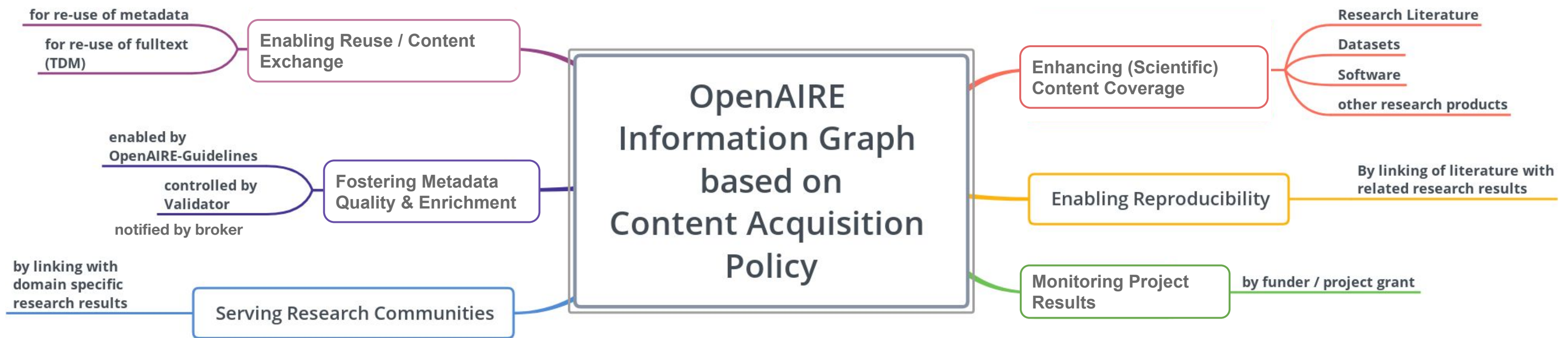


Agenda

- **OpenAIRE Aggregation Overview**
Objectives, Content, and Guidelines
- **Literature Guidelines v4 Overview**
Introduction to the Application Profile
- **Literature Guidelines v4 Implementation**
Repository Platforms, OpenAIRE Infrastructure, Pilot Data Sources
- **Conclusion / Last Words**

OpenAIRE Aggregation Overview

Objectives of OpenAIRE's Aggregation Policy



Content Acquisition Policy released 05-Oct-2018, <https://doi.org/10.5281/zenodo.1446408>

Integration Scenarios

- Directly harvested (from repositories, journals) and indirectly harvested (via aggregators, publishers)
- Individually claimed
- Individually referenced (if resolvable via PIDs)



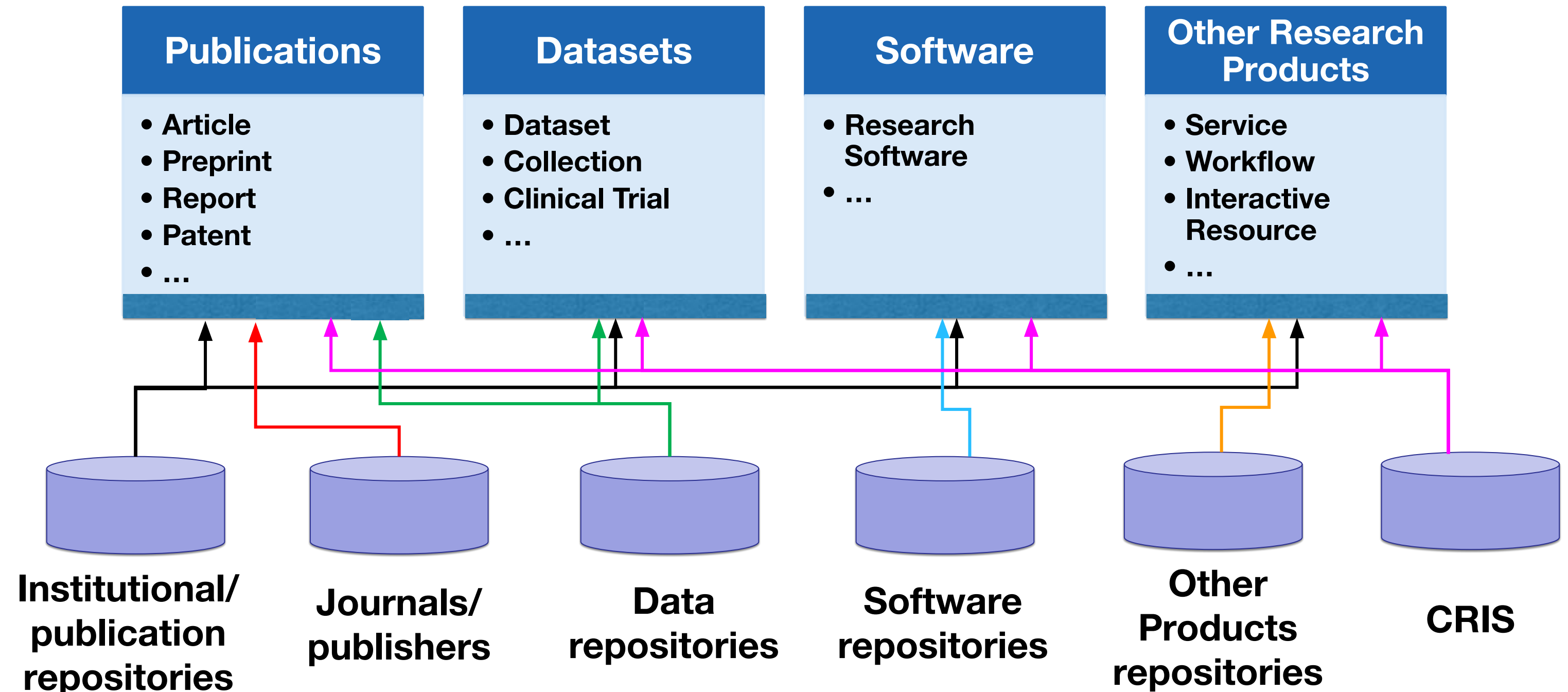
Link

Connect all your
research

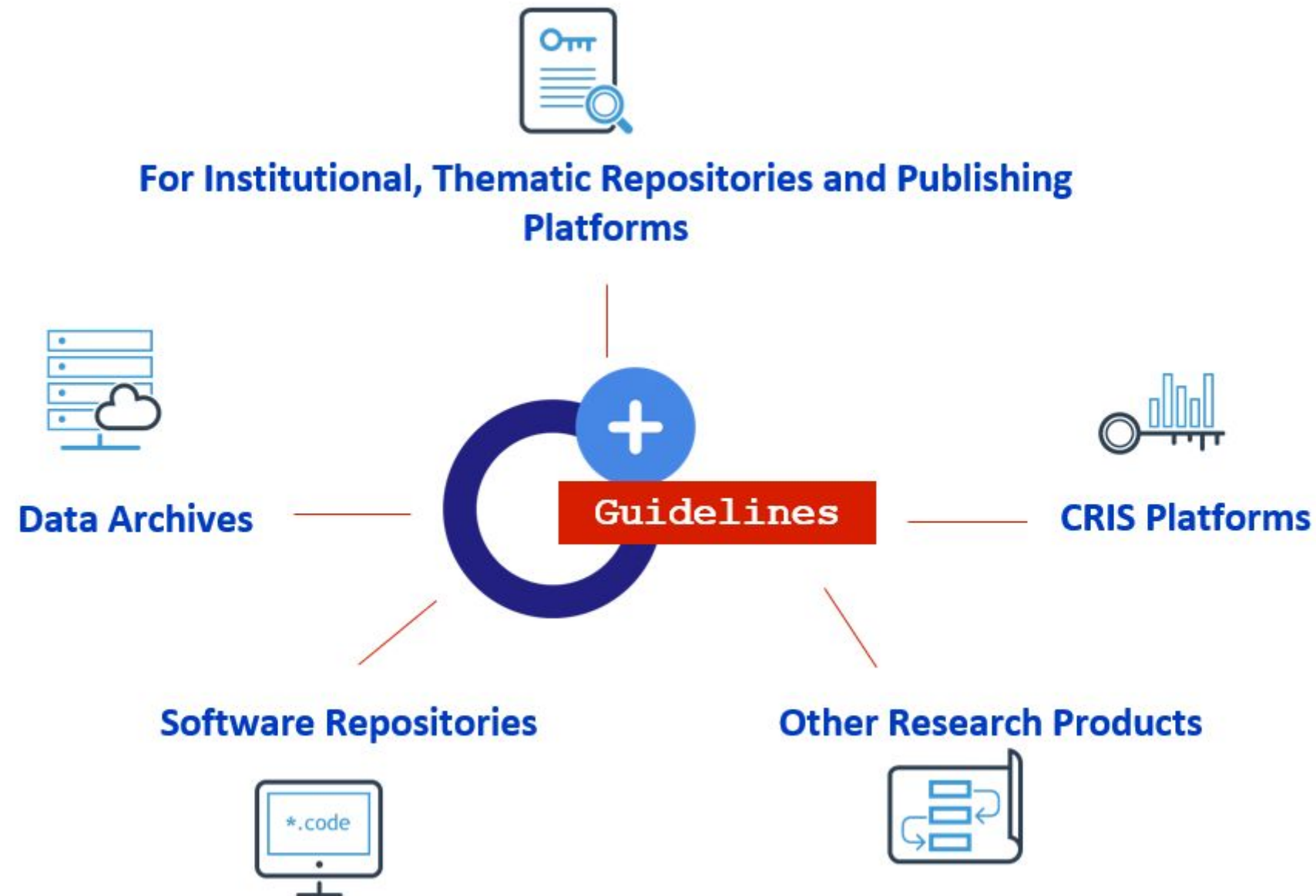
If you can't find your
research results in OpenAIRE,
don't worry!

Use our **LINK OUT service**,
that reaches out to many
external sources via APIs, to
connect your research results
and *claim* them to your
project.

Aggregated Artefact and Data Source Types



OpenAIRE's Guidelines for Open Science Content Providers



<https://guidelines.openaire.eu>

Guidelines v4

Overview

What's new / 1

- **Based on Dublin Core & DataCite metadata schemes**
- **Requires OAI metadata format prefix 'oai_openaire'**
- **Requires no specific OAI set**
 - Access level declaration per record more important
- **Compatible with new OpenAIRE Content Acquisition Policy**
 - Open access and non open access
 - Funded and non funded

What's new / 2

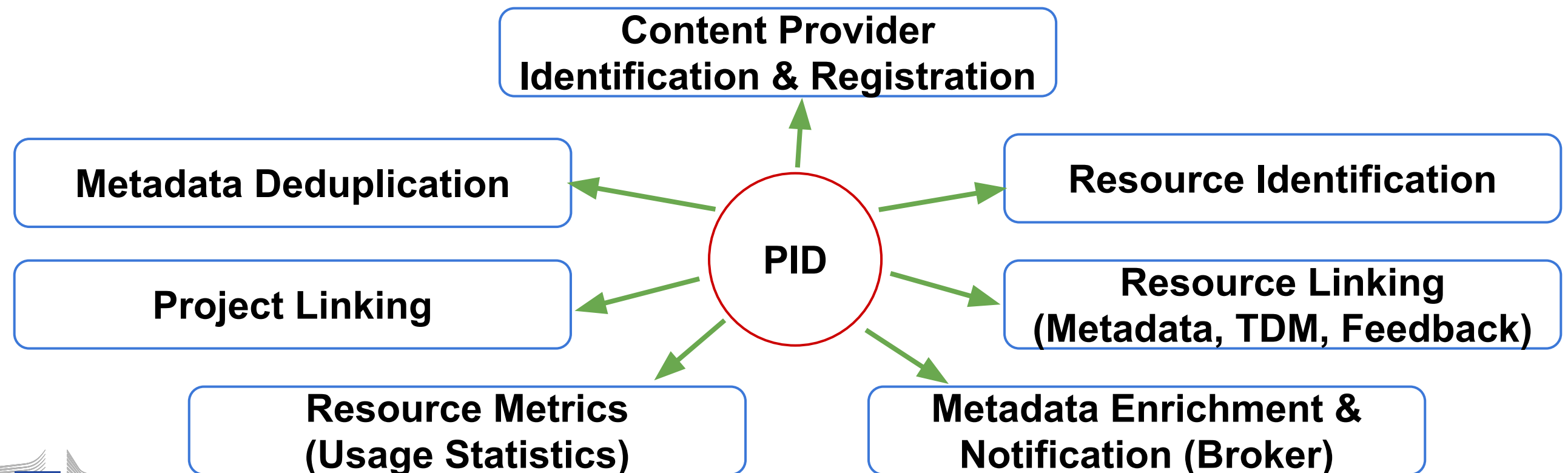
- **Uses controlled vocabularies defined by COAR**
(access rights, resource types, version types)

<https://www.coar-repositories.org/activities/repository-interoperability/coar-vocabularies/>

- **Enhances 'explicitness' by dedicated properties for specific information**

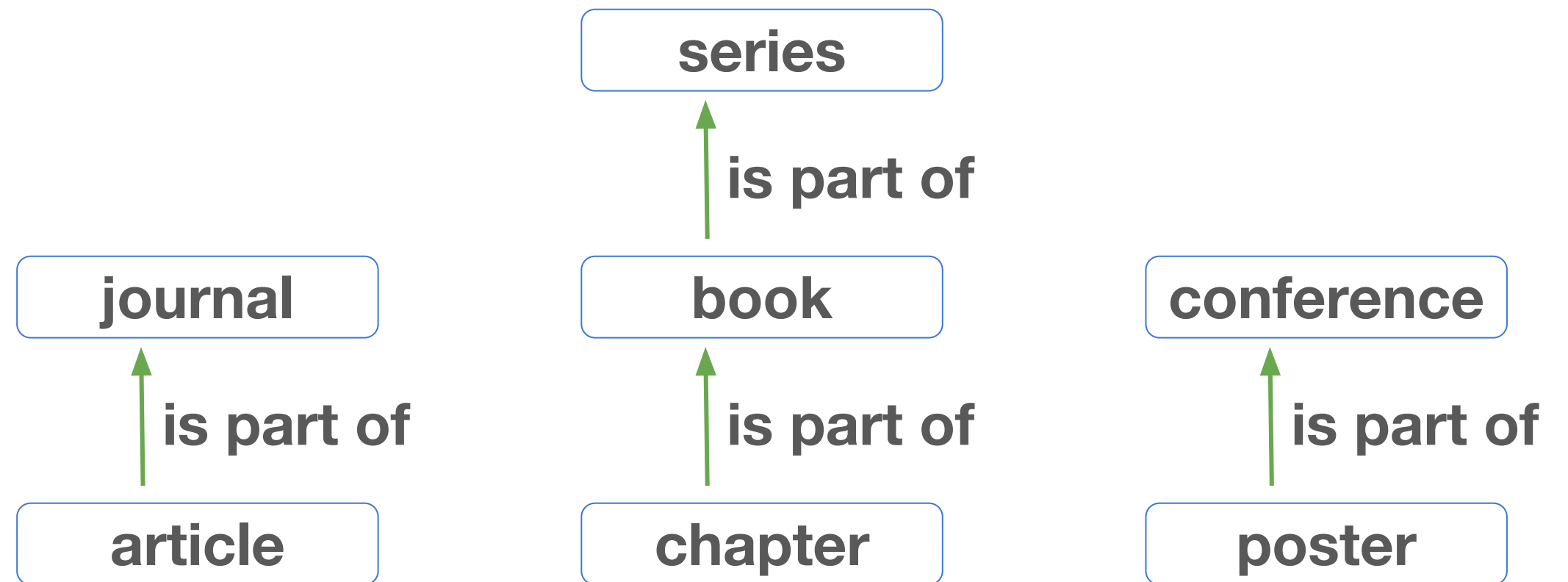
What's new / 3

- **Supports various identifier types**
 - ~ scholarly works: ARK, DOI, Handle, ...
 - ~ authors: ORCID, ISNI, ...
 - ~ organisations: ISNI, OpenDOAR, re3data, funder DOIs



What's new / 4

- Creates hierarchical level via container



OpenAIRE Guidelines for Literature Repositories v4.0

Application Profile

<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 | |

Metadata Goals

| Goal | Metadata Groups |
|--------------------------|--|
| discovery and citability | descriptive metadata |
| accessibility and reuse | access rights, licenses, file locations, dates (embargo end) |
| contextualization | research project, linked research artefacts |
| interoperability | entity identifiers, controlled vocabularies |
| reporting | funding references |

Discovery and Citability I

| Property | Requirement level | Granularity |
|---------------------|--------------------------------------|--|
| Creator | Mandatory | firstName, lastName, ID, person/organisation |
| Publication Date | Mandatory | |
| Title | Mandatory | language |
| Resource Identifier | Mandatory | identifier type |
| Resource Type | Mandatory / Controlled | resource category |
| Contributor | Mandatory if applicable | |
| Language | Mandatory if applicable / Controlled | |
| Description | Mandatory if applicable | language |
| Subject | Mandatory if applicable | language, classification/thesaurus |
| Publisher | Mandatory if applicable | |

Discovery and Citability II

| Property | Requirement level | Granularity |
|-------------------------------|-------------------|-------------|
| Source | Recommended | |
| Coverage | Recommended | |
| Audience | Optional | |
| GeoLocation | Optional | |
| Resource Version / Controlled | Recommended | |
| Citation Title | Recommended | |
| Citation Volume | Recommended | |
| Citation Issue | Recommended | |
| Citation Start Page | Recommended | |
| Citation End Page | Recommended | |
| Citation Edition | Recommended | |
| Citation Conference Place | Recommended | |
| Citation Conference Date | Recommended | |

hierarchical level via container
(for e.g. article-journal,
chapter-book, book-series,
presentation-conference)

Accessibility and Reuse

| Property | Requirement level | Granularity |
|---------------------|-------------------------|-------------|
| Access Rights | Mandatory / Controlled | |
| License Condition | Recommended | license URI |
| File Location | Recommended | |
| Embargo Period Date | Mandatory if applicable | |
| Size | Optional | |
| Format | Recommended | |

Contextualisation

| Property | Requirement level | Granularity |
|--------------------|-------------------------|--|
| Funding Reference | Mandatory if applicable | funderName, funderIdentifier, fundingStream, awardNumber |
| Related Identifier | Recommended | identifier type, relation type, resource type |

Interoperability

| Property | Requirement level | Granularity |
|---|--|--------------------------------|
| Resource Identifier | Mandatory | identifier type |
| Alternate Identifier | Recommended | identifier type |
| Related Identifier | Recommended | identifier type, relation type |
| Creator / nameIdentifier | Recommended | identifier type |
| Funding Reference / funderIdentifier / awardNumber | Recommended Mandatory if applicable | identifier type |
| | | |
| <i>data source identifier</i> | | |
| <i>Vocabularies</i> | Mandatory | |

Reporting (to Funders)

| Property | Requirement level | Granularity |
|-------------------|-------------------------|--|
| Funding Reference | Mandatory if applicable | funderName, funderIdentifier, fundingStream, awardNumber |

Guidelines References

- **Guidelines (OpenAIRE)**

<https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/>

- **Schema and examples (GitHub)**

<https://github.com/openaire/guidelines-literature-repositories>

- **Contribute / Feedback channels (various)**

- edit GoogleDoc: https://docs.google.com/document/d/1aFWrkBO_f_GTWWTAfco4uoUphEbtOOdJVD_2ohYw7Ek/edit
- visit GitHub: <https://github.com/openaire/guidelines-literature-repositories/issues>
- annotate guidelines: <https://web.hypothes.is>
- send emails: guidelines@openaire.eu
- contact NOADs: <https://www.openaire.eu/contact-noads>

Guidelines v4

Implementation

Implementation in Repository Platforms

| Software | Version | Status |
|------------------|--------------|---|
| DSpace | v7 v5, v6 | In preparation by DSpace OpenAIRE 4.0 WG Implementations by PT repos RCAAP for v5, in test |
| EPrints | all | Contacted |
| Invenio / zenodo | | On their roadmap |
| Islandora | | Contacted |
| LibreCat | | On their roadmap |
| OPUS | v4 | Contacted |
| MyCoRe | | Contacted |
| HAL | | Contacted |
| Haplo | version | |

Which further platforms need to be supported?

Implementation in OpenAIRE Infrastructure

| Service | Status |
|----------------------------------|--|
| Validator | in preparation |
| DataSource Registration & Update | ✓ |
| Aggregation | pre-implementation on beta in preparation |
| Indexing in explore portal | pre-implementation on beta in preparation |

If you'd like your Literature Guidelines v4 metadata integrated in OpenAIRE, please email to:

openaire-helpdesk@uni-bielefeld.de

Implementation in Pilot Data Sources

| Name | Organisation | Country | Platform |
|-----------------------------------|---------------------------|---------|----------|
| Arias Montano, Rep. Institucional | University of Huelva | Spain | DSpace |
| GREDOS | University of Salamanca | Spain | DSpace |
| Repositorio Documental | University of Valladolid | Spain | DSpace |
| ORBi | University of Liège | Belgium | DSpace |
| Refubium | Free University of Berlin | Germany | DSpace |
| Qeios | Qeios | UK | - |
| RADAR | Oxford Brookes University | UK | Equella |
| Rothamsted Research Rep. | Rothamsted Research | Uk | Haplo |

**Thanks for acting
as Literature v4
Guidelines pilots
and kind
cooperation!**

First Experiences with Metadata Quality

| Goal | Turn Attention to ... |
|---------------------------------|---|
| <i>metadata/XML reusability</i> | check correct namespaces, metadata root element, well-formed XML |
| discovery & citability | declare language for research items and metadata fields (title, description, subject) note applied subject classification/thesaurus mark creators as Personal or Organizational |
| accessibility & reusability | declare access rights and licenses |
| contextualization | declare/check relation to container entity (e.g. journal, book, series, conference): ~ relatedIdentifier with relationType IsPartOf for PIDs ~ citation* fields for title, volume, pages, etc. ~ start page and end page on 2 fields |
| interoperability | give all available PIDs for all entity types (research item, person, organisation, ...) |
| reporting | state/check explicit, pure funderName and/or funderIdentifier plus awardNumber (not just funder, not mingled together in shared fields) |

Conclusion / Last Words

Metadata Quality Aspects

- Timeliness: Metadata is up-to-date, i.e. records reflect the current state incl. recent changes.
- Completeness: All proper, suitable, relevant information is leveraged.
- Accuracy: Information is veritable, correct, non-contradictory. It is clear to which items statements refer (e.g. access rights or PIDs may refer to the metadata record itself, to the described item, to a cited item). Statements are sound, i.e. PIDs are valid.
- Legibility: Descriptions are easily comprehensible; due to language (e.g. titles), structure (of e.g. thesauri/vocabulary), homogeneity (e.g. order of author names).
- Consolidation: Records and statements are non-redundant, i.e. without duplicates. Records are contextualised, i.e. linked with other records (citations/references, versions/variants).
- Wastelessness: No test records, no records out of scope, no predatory sources, no spam, etc.
- Format-conformance: Compliance with format standards, utilisation of vocabularies/thesauri.

Conclusion / Last Words

- Upgrade to new Guidelines is recommended to improve metadata elaborateness
- Data sources following previous Guidelines versions will still be aggregated
- New Content Acquisition Policies (CAP) are valid
(all types of research output, OA and non-OA, funded and non-funded)
- A webinar on the new Guidelines will address repository managers
- Tell us which further platforms should be supported
- To get your Literature Guidelines v4 metadata integrated, email to openaire-helpdesk@uni-bielefeld.de
- Feedback on Guidelines is welcome (see feedback channel on earlier slide)

Thank you!

guidelines@openaire.eu

openaire-helpdesk@uni-bielefeld.de

<https://www.openaire.eu/contact-noads>

