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

OpenAIRE
Open REPOSITORIES 2019 | HAMBURG
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

http://box.openaire.eu/index.php/V9xSCykE5olkxMC
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

Content Providers
- Publications repositories
- Data repositories
- Registries
- Hybrid repositories

Research Infras
- PARThenos
- EPOS
- DARIAH-EU

TERMS OF USE

SchoolExplorer

Guidelines

Brokering
Cleaning
Validation
Harvesting

De-duplication
Inference
Uploading

Funding
Organization
Project
Source
Product
Publication
Data
Software
ORP

Publications
Repositories

Software
Repositories

DOAJ
Journals

ID

unpaywall
crossref
DOAJ
GRID

OpenAIRE Services
From basic infrastructure level to value added

- Research Analytics
- Institution Monitoring Dashboard
- Funder Monitoring Dashboard
- Open Science Observatory
- easyDMP
- Anonymization
- OpenAIRE CONNECT
- OA Broker
- Interoperability Guidelines
- APIs
- Validator
- Usage Analytics
- OpenAIRE EXPLORE
- OpenAIRE PROVIDE
- OpenAIRE DEVELOP
- Zenodo
26 million publications, 1 million research data from 15 thousand content providers and 18 funders linked together for an integrated research search.

94 million publications, 8 million research data from 9,800 content providers and 28 funders linked together for an integrated research search.
OpenAIRE graph expansion

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

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 Access

Open Science

Open Science Monitor

Updated Methodological Note

Brussels, October 8th 2018

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

Subcontractor: Elsevier
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 +
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: 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)
Open metadata

Repositories
Archives

Repositories
Archives

Springer
Microsoft
De-duplicated

<table>
<thead>
<tr>
<th>Entity type</th>
<th># Collected records</th>
<th># Records after cleaning and de-duplication</th>
<th># Identified duplicates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publications</td>
<td>~ 343M</td>
<td>~ 94M</td>
<td>~ 249 millions</td>
</tr>
<tr>
<td>Data</td>
<td>~ 5.2M</td>
<td>~4.6M</td>
<td>~ 600K</td>
</tr>
<tr>
<td>Software</td>
<td>~150K</td>
<td>~ 134K</td>
<td>~ 20K</td>
</tr>
<tr>
<td>Other</td>
<td>~ 5M</td>
<td>~ 4.5M</td>
<td>~ 500K</td>
</tr>
<tr>
<td>Organisations</td>
<td>~ 380K</td>
<td>~220K</td>
<td>~ 160K</td>
</tr>
</tbody>
</table>

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](provide.openaire.eu))

- Openly accessible via APIs ([develop.openaire.eu](develop.openaire.eu))
Trusted (in progress)

- Authors in the loop to enrich their ORCID record
- Validation of end-user "claims"
Monitoring Tools
OpenAIRE Open Science Monitoring

Research impact

- MONITOR
- CONNECT

Open Access/Science trends

- MONITOR
- CONNECT

Research admins

Funders

Research communities

Research Organization
CONNECT
Research Community Dashboard for Research Infrastructures (Initiatives)
For building community-specific gateways to Open Science

A virtual environment to implement Open Science publishing practices

Uptake of Open Science publishing practices

MONITOR

DEPOSIT ANYTHING

Research Impact
Report to funders

Research Community Dashboard

CONNECT

All the relevant research products
… linked

On demand publishing on services of Research Infrastructures
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

<table>
<thead>
<tr>
<th>Funding Impact</th>
<th>Open Access/Science Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publications, research data, software published thanks to the existence of the RI</td>
<td>Monitoring of Open Science impact: data/software FAIRness, reproducibility trends</td>
</tr>
</tbody>
</table>
Added value functionalities

• Open Science indicators
MONITOR

Funder Dashboard
(Project dashboard)
<table>
<thead>
<tr>
<th>Publications</th>
<th>Datasets</th>
<th>Software</th>
<th>Funders</th>
</tr>
</thead>
<tbody>
<tr>
<td>25412485</td>
<td>1017048</td>
<td>93659</td>
<td>18</td>
</tr>
</tbody>
</table>

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

**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 (MSEŠ)
- 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
  - MESD - Ministry of Education, Science and Technological Development of Republic of Serbia
- Switzerland
  - SNSF - Swiss National Science Foundation
- Turkey
  - TÜBİTAK - Türkiye Bilişim ve Teknolojik Araştırma Kurumu
- United Kingdom
  - RCUK - Research Council UK
  - WT - Welcome Trust

**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
## Aims

<table>
<thead>
<tr>
<th>Funding Impact</th>
<th>Open Access/Science Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publications, research data, software published thanks to grants awarded by the funder</td>
<td>Monitoring of Open Science impact: data/software FAIRness, reproducibility trends</td>
</tr>
</tbody>
</table>
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 mentioned? Tell us!

Explore the BETA graph and tell us how to improve!

https://beta.explore.openaire.eu
Questions?
OpenAIRE content acquisition policy

new terms of agreement for content providers
ALL Literature, Research data, Software, Other research products

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

www.openaire.eu/content-acquisition-policy

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
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.

how we process
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
“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

**Coverage**
OpenAIRE will actively pursue harvesting content from European but also non-European repositories.

**Reproducibility**
The OpenAIRE graph aims at linking scientific literature with all products used or resulting from the related research activity.

**Research Communities**
The OpenAIRE graph links research products with communities for which they are relevant, in order to provide a multi-community view of the scholarly output.

**Monitoring**
The OpenAIRE graph links research products with the tenders and projects resulting from their grants.

**Quality**
Data sources and repositories are quality-controlled: their metadata respect the OpenAIRE guidelines, and their import is curated by data curators.

**Terms of Agreement**
Content providers, managers, and users accept the OpenAIRE Terms of Agreement in order for OpenAIRE to re-use their metadata and OA full-text under specific terms, warranties and license.
Objectives of OpenAIRE’s Aggregation Policy

OpenAIRE Information Graph based on Content Acquisition Policy

- For re-use of metadata
- For re-use of fulltext (TDM)
- Content Quality
- Content Coverage (according to OpenAIRE-Guidelines)
- Enabling Reproducibility
- Monitoring

- Serving Research Communities
- Research Literature
  - Datasets
  - Software
  - Other research products
- By linking of literature with related research results

Enabled by OpenAIRE-Guidelines
Controlled by Validator
By linking with domain specific research results

by funder / project grant
CAP supported by the set of Guidelines for Open Science Content Providers

https://guidelines.openaire.eu
26 mi publications, 1 mi research data from 15 K content providers and 18 funders linked together for an integrated research search.

https://beta.explore.openaire.eu

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

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
Novelties: orcid, collected from, communities...
New data sources

- EGI Application database
- OMICS DI
- Kaggle
- ReactToMe
- DOECODE
- Unpaywall
## Metadata Quality Challenges

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

• Document classification:
  https://explore.openaire.eu/search/publication?articleId=od______1874::6331f80a2b9758f56609a874e9adc26

• 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=od________18::4405ffbf18cc37d73d0daff3650e48f82](https://explore.openaire.eu/search/publication?articleId=od________18::4405ffbf18cc37d73d0daff3650e48f82)

- Link from a software to its «main» publication: [https://explore.openaire.eu/search/software?softwareId=openaire::949d7264f0efb7a27e521fee9c59209b](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](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
- Link to project: https://explore.openaire.eu/search/publication?articleId=dedup_wf_001::4be652d611c4bbc897118bdb564c557
Quality issues

• Take care of your PIDs:
  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
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
Evolution of OpenAIRE-Guidelines

2010
Literature Guidelines v1

2012
- Literature Guidelines v2
- Data Guidelines v1

2013
Literature Guidelines v3

2014
Data Guidelines v2

2015
CRIS-CERIF Guidelines v1

2018 Guidelines for
- Software Repositories
- Other Research Products

2018 Guidelines for
- institutional and thematic repos. v4.0
- CRIS-CERIF v1.1
Diversity of Research Results from Different Types of Sources

**Publications**
- Article
- Preprint
- Report
- ...

**Datasets**
- Dataset
- Collection
- Clinical Trials
- ...

**Software**
- Research Software
- ...

**Other Research Products**
- Service
- Workflow
- Interactive Resource
- ...

Repositories:
- Institutional/publication repositories
- Journals/publishers
- Data repositories
- Software repositories
- Other Products repositories
<table>
<thead>
<tr>
<th>Goal</th>
<th>Metadata Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery and Citability</td>
<td>Descriptive metadata</td>
</tr>
<tr>
<td>Accessibility and Reuse</td>
<td>Access Rights, License Conditions</td>
</tr>
<tr>
<td>Contextualization</td>
<td>Research Project, Linked Research Artefacts</td>
</tr>
<tr>
<td>Interoperability</td>
<td>Identifier for Entities, Controlled Vocabularies</td>
</tr>
<tr>
<td>Reporting</td>
<td>Funding Reference</td>
</tr>
<tr>
<td>TDM</td>
<td>File Location, License Conditions</td>
</tr>
</tbody>
</table>
OpenAIRE’s Guidelines for Open Science Content Providers

For Institutional, Thematic Repositories and Publishing Platforms

Data Archives

Software Repositories

CRIS Platforms

Other Research Products

https://guidelines.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.
Role of PIDs in OpenAIRE

- Content Provider Identification & Registration
- Resource Identification
- Resource Linking (by Inference & User-Feedback)
- Metadata Deduplication
- Metadata Enrichment & Notification (Broker Service)
- Resource Metrics (Usage Statistics)
OpenAIRE Guidelines for Literature Repository Managers v4.0

• Established standards: [Dublin Core](https://dublincore.org) and [DataCite](https://datacite.org) 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)
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

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

• Need for a specific format for scientific publications
• More specific fields for metadata fields
• Support for hierarchichal information (concept of entities)
• Metadata Alignement 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 respectivo identificador no ORCID ou Ciência Vitae.

- Expose in OpenAIRE 4 metadata schema (OAI-PMH)
Transformation of project ID into project entity

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

```xml
<fundingReferences>
  <fundingReference>
    <funderName>European Commission</funderName>
    <funderIdentifier fundIdentifierType="Crossref Funder ID">http://doi.org/10.13039/5011000008500</funderIdentifier>
    <fundingStream>FP7</fundingStream>
    <awardNumber>612425</awardNumber>
  </fundingReference>
</fundingReferences>
```
Submission Forms (authors)

- Author IDs
- COAR Taxonomies
Expose OAI-PMH with oai_openaire

```xml
  <dc:titles>
    <dc:title xml:lang="pt-PT">Embargoed Item Title</dc:title>
  </dc:titles>
  <dc:creators>
    <dc:creator>
      <dc:creatorName>Carvalho, José</dc:creatorName>
      <dc:givenName>José</dc:givenName>
    </dc:creator>
    <dc:creator>
      <dc:creatorName>Carvalho</dc:creatorName>
      <dc:familyName>Carvalho</dc:familyName>
    </dc:creator>
  </dc:creators>
  <dc:contributors>
    <dc:contributor contributorType="HostingInstitution">
      <dc:contributorName>DSpace AuthorProfile</dc:contributorName>
    </dc:contributor>
    <dc:contributor>
      <dc:nameIdentifier nameIdentifierScheme="orcid">0000-0003-1729-3404</dc:nameIdentifier>
    </dc:contributor>
    <dc:nameIdentifier nameIdentifierScheme="cienciad">6013-8C52-C280</dc:nameIdentifier>
  </dc:contributors>
  <fundingReferences>
    <fundingReference>
      <funderName>European Commission</funderName>
      <funderIdentifier funderIdentifierType="Crossref Funder ID">http://doi.org/10.13039/501100008530</funderIdentifier>
      <fundingStream>FP7</fundingStream>
    </dc:fundingReference>
  </fundingReferences>
</resource>
```
What Changes? On Journals (OJS)...

- OAI PMH (not yet developed)
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
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;
Alessia Bardi, CNR-ISTI, alessia.bardi@isti.cnr.it
André Vieira, University of Minho, andrevieira@sdum.uminho.pt;
Jochen Schirrwagen, Bielefeld University, jochen.schirrwagen@uni-bielefeld.deinfo@openaire.eu
info@openaire.eu