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

Pedro Príncipe, University of Minho, André Vieira, University of Minho, Alessia Bardi, CNR-ISTI, Jochen Schirrwagen, Bielefeld University, Dimitris Pierrakos, ATHENA RC
AGENDA

13:30 – Welcome and introduction, Pedro Príncipe
13:40 – OpenAIRE catch-all Broker Service for literature and data repositories
13:55 – Dashboard for Content Providers: Open Science as-a-Service for repositories
14:10 – Dashboard demo and test drive (breakout groups)
15:00-15:30 – Coffee break
15:30 – OpenAIRE usage statistics service - overview of the service
15:50 – Service demo base on real case
16:10 – Collection workflows
16:30 – How to join OpenAIRE usage statistics, Tracking usage events
16:45 – Discussion
17:00 – Closing
Workshops Topics – 2nd part

1) OpenAIRE catch-all Broker Service for literature and data repositories
2) Dashboard for Content Providers: Open Science as-a-Service for repositories

>>> 14:10 – Dashboard demo and test drive (breakout groups)

3) OpenAIRE usage statistics service - overview of the service
4) Service demo base on real case & Collection workflows
5) How to join OpenAIRE usage statistics, Tracking usage events

>>> Discussion
SLIDES HERE:

bit.ly/openaire_or2019

http://box.openaire.eu/index.php/s/V9xSOyKE5okbxMC
1

Broker Service

Catch-all broker for content providers
Repositories in OpenAIRE may be interested to acquire metadata information about publications that are “potentially of interest to them” i.e. be part of their collection: add new records, enrich the records with extra metadata information.
OpenAIRE Broker sketch

Identifying “events” relevant to repositories (enrichments & additions)

Event (potential notification):
- Message
- Topic
- TargetRepository
- Trust

OpenAIRE Information Space Graph
(deduplication, Inference, Aggregation)

OpenAIRE Data Sources
Enrichments events from the broker service

- **ENRICH/MISSING/PROJECT** - OpenAIRE discovered references to research projects that can be associated to your publications
- **ENRICH/MISSING/OPENACCESS_VERSION** - OpenAIRE discovered Open Access versions of your publications
- **ENRICH/MISSING/PID** - OpenAIRE discovered missing persistent identifiers associated to your publications
- **ENRICH/MISSING/ABSTRACT** - OpenAIRE discovered missing abstracts among your publications
- **ENRICH/MISSING/DATE** - OpenAIRE discovered missing date values to your publications
- **ENRICH/MISSING/SUBJECT/MESHEUROPMC** - OpenAIRE discovered classification terms from the Medical Subject Headings that can be associated to your publications
- **ENRICH/MISSING/SUBJECT/JEL** - OpenAIRE discovered Journal of Economic Literature (JEL) classification terms that can be associated to your publications
- **ENRICH/MISSING/SUBJECT/DDC** - OpenAIRE discovered Dewey Decimal classification terms (DDC) that can be associated to your publications
- **ENRICH/MISSING/SUBJECT/ACM** - OpenAIRE discovered ACM classification terms that can be associated to your publications
- **ENRICH/MISSING/SUBJECT/ARXIV** - OpenAIRE discovered ARXIV classification terms that can be associated to your publications
Enrichments:

- MORE metadata
- MISSING metadata
Dashboard for Content Providers

Service demo and test drive: content enrichment events.
One-stop-shop web service where content providers (repositories, data archives, journals, aggregators, CRIS systems) interact with OpenAIRE. It provides the front-end access to many of OpenAIRE's backend services.
Content Provider Dashboard

A one-stop-shop for sharing, finding and enriching your content

START HERE

Validator
Interoperability Guidelines
OA Broker
Usage Analytics

Validate
Interoperable metadata is key for effective content sharing.

Register
Reach a wider audience around the world

Enrich
Improve your metadata. Get more connections

Measure
Open research impact empowers Open Science
one stop shop for all OpenAIRE content providers

Services for repository managers, Data Archive managers, OA publishers...

SUPPORT, PROMOTE, ENRICH
## Dashboard usage

<table>
<thead>
<tr>
<th>TARGET USERS</th>
<th>USER BASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repository managers (literature, data), libraries, content providers, publishers, national aggregators.</td>
<td>800 content providers have used the registration and validation service (<em>V1 focused on literature repositories, 1200 all types for V2</em>).</td>
</tr>
</tbody>
</table>

### USER VALUE

*Improve repository collections* and content for *enhanced visibility* and access. Improved institution memory. Better institution research assessment. *Compliance to funder rules*. Improved repository *interoperability*. 
OpenAIRE Content Provider Dashboard

what it does
Content Provider Dashboard

Register
Register data sources in the OpenAIRE infrastructure

Validate
Validate data sources against OpenAIRE guidelines

Notifications
View notifications to enrich the metadata and the content

Metrics
View aggregated, cleaned usage statistics for repository access
Interoperable metadata is key for effective content sharing

Use our validation service and see how you can apply the OpenAIRE Guidelines to expose your contents using global standards.
Reach a wider audience around the world
Register your datasource in OpenAIRE and be part of a global interlinked network.
Improve your metadata. Get more connections

OA Broker service offers a wealth of information on scholarly communication data.

Find out what interests you and subscribe to enrich your records.

More & Missing events that may enrich your Repository:

- Persistent identifiers
- Open Access Versions
- Projects
- Subjects
- Abstracts
  … datasets, software
Open research impact empowers Open Science

Open Metrics service by sharing your usage data. Get the benefit of an aggregated environment to broaden the mechanisms for impact assessment.

Get usage statistics reports for your datasource

- **AR1**: Article Report 1, number of successful article download requests by month and repository.
- **IR1**: Item Report 1, number of successful item download requests by month and repository.
- **RR1**: Repository Report 1, number of successful item downloads for all repositories participating in the usage statistics service.
- **BR1**: Book Report 1, number of successful title requests by month and title.
- **BR2**: Book Report 2, number of successful section requests by month and title.
Novelties
Enrichments for Universidade do Minho: Repository

MORE

<table>
<thead>
<tr>
<th>EVENT</th>
<th># OF EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENRICH/MORE/OPENACCESS_VERSION</td>
<td>18296</td>
</tr>
<tr>
<td>OpenAIRE discovered another Open Access version of a publication</td>
<td></td>
</tr>
<tr>
<td>ENRICH/MORE/PID</td>
<td>5725</td>
</tr>
<tr>
<td>OpenAIRE discovered another persistent identifier associated to your publications</td>
<td></td>
</tr>
<tr>
<td>ENRICH/MORE/SUBJECT/JEL</td>
<td>4154</td>
</tr>
<tr>
<td>OpenAIRE discovered more Journal of Economic Literature (JEL) classification terms that can be associated to your publications</td>
<td></td>
</tr>
</tbody>
</table>

MISSING

<table>
<thead>
<tr>
<th>EVENT</th>
<th># OF EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENRICH/MISSING/PROJECT</td>
<td>3422</td>
</tr>
<tr>
<td>OpenAIRE discovered references to research projects that can be associated to your publications</td>
<td></td>
</tr>
<tr>
<td>ENRICH/MISSING/PID</td>
<td>3132</td>
</tr>
<tr>
<td>ENRICH/MISSING/SUBJECT/DDC</td>
<td>736</td>
</tr>
<tr>
<td>OpenAIRE discovered Dewey Decimal classification terms (DDC) that can be associated to your publications</td>
<td></td>
</tr>
</tbody>
</table>

METADATA ENRICHMENTS FROM THE OPENAIRE BROKER SERVICE (monthly notifications)
Collection monitor provides the aggregation history.
THE SERVICE COLLECTS AND EXPLOITS USAGE METRICS LIKE DOWNLOADS AND METADATA VIEWS (tracking plugin, SUSHI endpoint, COUNTER reports)
provide.openaire.eu

Content provider?
Join OpenAIRE, use our tools and make your content more visible around the world.

OpenAIRE.PROVIDE

In production since October!
Content Provider Dashboard: testing phase

**Portugal**
- 1 Webinar, 59 Attendees
- 21 Repositories represented

**Spain**
- 1 Webinar, 49 Attendees
- 28 Repositories represented

**Real users of the Dashboard**
- Repository Managers from Portugal & Spain

**2 Webinars (Demo)**
- Main OpenAIRE services for content providers (Dashboard and Broker), RCAAP and RECOLECTA.

**Test drive**
- Grant the access to the Dashboard
- Guidance on the functionalities usage

**Collect feedback**
- Questionnaire & Helpdesk

May/June 2018
Content Provider Dashboard

A one-stop-shop for sharing, finding and enriching your content

Start here

provide.openaire.eu
Breakout groups – demo
Login: repositiorium
Pass: openaire1!

Login: comum
Pass: Rcaapsdum2008
Support materials for Content Providers Dashboard uptake

Support – guides

• Provide - How to validate and register your repository
• Provide - How to enrich research artifacts
• Usage Statistics – How to track the usage activity of your repository
• ScholExplorer - Literature & Data interlinking
• Making your repository Open

Training – webinars

• Make your content count - OpenAIRE Content providers Dashboard: service for repository managers
• OpenAIRE metrics service: usage statistics
• OpenAIRE Guidelines for data providers: new Metadata Application Profile for Literature Repositories
what comes next
New broker events

- ORCID > Publications/Datasets/software
- Datasets > Projects
- Publications <-> Software
- Publications <-> Datasets

<table>
<thead>
<tr>
<th>Event Description</th>
<th># of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>EnrichMore/OpenAccess/Version</td>
<td>20029</td>
</tr>
<tr>
<td>EnrichMore/PID</td>
<td>3208</td>
</tr>
<tr>
<td>EnrichMore/Subject/Arxiv</td>
<td>1575</td>
</tr>
<tr>
<td>EnrichMore/Subject/Jel</td>
<td>1257</td>
</tr>
<tr>
<td>EnrichMore/Subject/Doc</td>
<td>9713</td>
</tr>
<tr>
<td>EnrichMore/Subject/MeshBrief/Medline</td>
<td>757</td>
</tr>
<tr>
<td>EnrichMore/Subject/ACM</td>
<td>145</td>
</tr>
<tr>
<td>Total</td>
<td>28034</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event Description</th>
<th># of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>EnrichMissing/Project</td>
<td>8735</td>
</tr>
<tr>
<td>EnrichMissing/OpenAccess/Version</td>
<td>1794</td>
</tr>
<tr>
<td>EnrichMissing/PID</td>
<td>726</td>
</tr>
<tr>
<td>EnrichMissing/Abstract</td>
<td>146</td>
</tr>
<tr>
<td>EnrichMissing/Subject/MeshBrief/Medline</td>
<td>87</td>
</tr>
<tr>
<td>EnrichMissing/Subject/Jel</td>
<td>44</td>
</tr>
<tr>
<td>EnrichMissing/Subject/Doc</td>
<td>23</td>
</tr>
<tr>
<td>EnrichMissing/Subject/ACM</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>11545</td>
</tr>
</tbody>
</table>
www.openaire.eu
@openaire_eu
facebook.com/groups/openaire

pedroprincipe@sdum.uminho.pt
OpenAIRE USER FEEDBACK FORM: bit.ly/brokerevaluation

https://forms.gle/uQZQ8954g47az7Vn8
3

OpenAIRE usage statistics service

Overview, demo and collection workflows
Why Usage Statistics

- Authors / readers: to assess impact and to compare with scholarly works of peers
- Repository managers: how often / how well is the repository service used
- Research funders / administrators: usage metrics as part of a variety of metrics to assess impact of research funded or supported

Open research impact empowers Open Science
How it works

- (1) Tracking of usage events or (2) retrieving COUNTER reports on scholarly work items in repositories
- Applying rules based on COUNTER CoP
- Aggregating metrics on usage for scholarly works
- Repository plugins to track usage events
TIERS Collection Workflows for Usage Statistics
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>idSite</td>
<td>the ID of the repository</td>
</tr>
<tr>
<td>idVisit</td>
<td>a visitor/session ID (an 8 byte binary string)</td>
</tr>
<tr>
<td>visitIP (optionally anonymized)</td>
<td>the IP address of the visitor</td>
</tr>
<tr>
<td>action</td>
<td>the action performed (view, download, outlink, etc)</td>
</tr>
<tr>
<td>url</td>
<td>the url of the requested item</td>
</tr>
<tr>
<td>timestamp</td>
<td>the date &amp; time of the request</td>
</tr>
<tr>
<td>OAI-PMH Identifier</td>
<td>the Open Access Initiative identifier of the item being viewed/downloaded</td>
</tr>
<tr>
<td>agent</td>
<td>the Web Browser and the operating system of the visitor</td>
</tr>
<tr>
<td>referrer</td>
<td>The url linked to the item requested</td>
</tr>
</tbody>
</table>
Usage Statistics on Datasource and Item Level
OpenAIRE Exposes Usage Statistics

http://services.openaire.eu/usagestats/sushilite/

COUNTER compliant usage statistics reports for Repositories

AR1
Article Report 1, number of successful article download requests by month and repository.

IR1
Item Report 1, number of successful item download requests by month and repository.

RR1
Repository Report 1, number of successful item downloads for all repositories participating in the usage statistics service.

BR1
Book Report 1, number of successful title requests by month and title.

BR2
Book Report 2, number of successful section requests by month and title.
SUSHI response example (JSON)

Repository Report

```json
- { ItemIdentifier: []
  - { Type: "OpenAIRE",
      Value: "openiar___:8e98d81f8217304975cc23337bb5761"
    }
  - { Type: "OpenDOAR",
      Value: "307"
    }
  - { Type: "URL",
      Value: "https://repositorium.sdum.uminho.pt/"
    }
  ItemPlatform: "Universidade do Minho: RepositoriUM",
  ItemDataType: "Platform",
  ItemPerformance: []
  - { Period: {
      Begin: "2017-01-01",
      End: "2017-01-31"
    }
  },
  Instance: []
  - { MetricType: "fr_total",
      Count: "22087"
    },
  - { MetricType: "abstract",
      Count: "51685"
    }
  Category: "Requests"
}
```

Item Report

```json
Report: {
  @Created: "2017-09-06 08:08:21+0000",
  @Version: "4",
  @Name: "IRI:4",
  Vendor: {
    Contact: {
      Contact: "OpenAIRE HelpDesk",
      E-mail: "helpdesk@openaire.eu"
    }
  },
  Name: "OpenAIRE"
},
Customer: {
  ID: "anonymous",
  ReportItems: []
  - { ItemIdentifier: []
    - { Type: "OpenAIRE",
        Value: "033328d82037702785e0895601849ef"
      }
    - { Type: "URL",
        Value: "http://hdl.handle.net/1822/7975 ; http://hdl.handle.net/1822/7465 ; http://europepmc.org/articles/P1K2008319 ;"
      }
    - { Type: "DOI",
        Value: "10.1007/s10879-017-0425-4"
      }
    },
    ItemPublisher: "American Society for Microbiology",
    ItemPlatform: "Universidade do Minho: RepositoriUM",
    ItemDataType: "Article",
    ItemName: "Adaptive evolution of a lactose-consuming Saccharomyces cerevisiae recombinant",
    ItemPerformance: []
    - { Period: {
        Begin: "2017-01-01",
        End: "2017-01-31"
      }
    },
    Instance: []
    - { MetricType: "ft_total",
        Count: "5"
      },
    - { MetricType: "abstract",
        Count: "4"
      }
    },
    Category: "Requests"
  }
}
Main Challenges and Limitations

• **Comparability of Usage Statistics across platforms**
  • Must be collected and processed by agreed common standards, e.g. COUNTER CoP
  • Must be openly accessible by default (e.g. CC-0 license)

• **Comparability of Usage Statistics of different versions of an item**
  • e.g. Usage Statistics of Open Access vs. non-Open Access item versions
  • Usage Statistics must be enriched by item identifiers
  • Items must have comprehensive metadata descriptions

• **Contextualized Usage Statistics**
  • COUNTER Reports are a basic first step but limited on statistics per platform (items)
  • Aggregated usage statistics of deduplicated items
  • Linked research results and their Usage Statistics in the context of (e.g. a project or topic)

"The larger the network the better the usage statistics"
Next Steps

- Promote the service to content provider managers
- Update to COUNTER CoP R5
- Support of COUNTER CoP for Research Data (work with MakeDataCount)
- Improve visualizations of Usage Statistics in portal and dashboard
- Aligning tracking protocols and generic tracking solutions
- Support national usage statistics initiatives to become a node in OpenAIRE Usage Statistics

collaboration already with LA Referencia, IRUS-UK
More Information

- Item in the EOSC Service catalogue: https://catalogue.eosc-portal.eu/service/openaire.openaire_usage_statistics
- OpenAIRE2020 - Usage Statistics Services - D8.5: https://doi.org/10.5281/zenodo.1034164
Architecture and Workflows for Usage Statistics

PUSH
tracked event

Matomo
processing script

Metadata-Index

UsageStatistics-DB

PULL
COUNTER Report

processing script

Repository
- CRIS
- eJournal

National Statistics Node
PUBLISHER
Metrics in the OpenAIRE Dashboard for content providers

- Four steps to join OpenAIRE Usage Statistics
  1. Download.
  2. Configure.
  3. Deploy.
  4. Validate (by OpenAIRE).

- Or enter SUSHI endpoint to let OpenAIRE collect COUNTER reports
Enable Metrics for selected Datasource

YOU DON'T HAVE METRICS ENABLED FOR THIS REPOSITORY YET. WOULD YOU LIKE TO ENABLE THEM?

Once you select to enable metrics for your repository, the following steps need to be performed:

**On your side**
1. Download the tracking code for your repository platform
2. Configure the tracking code according to the instructions
3. Deploy the tracking code in your repository platform

**On the OpenAIRE's side**
4. Validate the installation of the tracking code and inform the repository manager accordingly

For more details about the workflows and tools please consult the "Guidelines for Collecting Usage Events and Provision of Usage Statistics".
Configure Metrics for selected Datasource

Metrics Configuration & Software Details

OpenAIRE’s usage statistic service uses the Piwik Open Source Analytics platform (piwik.org) to track usage activity. When metrics are enabled for a repository, two unique identifiers are generated - a piwik-ID that associates the repository with its usage events in Piwik and an authentication-ID that allows to track usage activity on the Piwik platform. Metadata views and item downloads are tracked and automatically sent to Piwik. Statistics are generated using the COUNTER Code of practice directives.

OpenAIRE’s usage statistics service tracking code exploits Piwik’s API. In order to make the tracking of usage events from repositories more robust, it was necessary to implement repository platform specific patches and plugins starting with DSpace and EPrints. The code is maintained on Github:

- as a patch for various versions of DSpace (https://github.com/openaire/OpenAIRE-Piwik-DSpace)

To configure your repository to allow tracking in Piwik platform, please change the configuration files with the following parameters and values, generated for your site:

PiwikID
000

AuthenticationToken
01233456

Details for the configuration files are given in the README of the tracking code.
Matomo Live tracking (Pull)

- Demo Site:
  
  http://demo.librecat.org/

- Real time tracking:
  
  https://analytics.openaire.eu/index.php?module=CoreHome&action=index&idSite=1&period=day&date=yesterday&updated=1#?idSite=1&period=day&date=yesterday&category=Dashboard_Dashboard&subcategory=1&popover=
Matomo offline collection of COUNTER Reports (Push)

- Irus-UK (aggregator) endpoint via SUSHI-Lite:
  
  https://irus.jisc.ac.uk/api/sushilite/v1_7/Client/IR1/

- Aberdeen University Research Archive (irusuk:28)

  https://irus.jisc.ac.uk/api/sushilite/v1_7/GetReport/?Report=IR1&Release=4&RequestorID=Jisc&BeginDate=2019-01&EndDate=2019-03&RepositoryIdentifier=irusuk%3A28&ItemIdentifier=&ItemType=&hasDOI=&Granularity=Monthly&Callback=&Pretty=Pretty
Usage Statistics Reports

- Repository Statistics from the Content Providers Dashboard

provide.openaire.eu

- COUNTER reports using the SushiLite API

http://services.openaire.eu/usagereports/sushilite/

http://services.openaire.eu/usagereports/sushilite/GetReport/?Report=IR1&Release=4&RequestorID=anonymous&BeginDate=2019-01&EndDate=2019-02&RepositoryIdentifier=opendoar%3A1714&ItemIdentifier=&ItemDataType=&Granularity=Monthly&Pretty=Pretty
Questions

- How important the Metrics service is for your repository?
- Do you consider Usage statistics as a “trustworthy” evaluator for scholarly impact?
www.menti.com

192517
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