

# EVENTOS

XVI WORKSHOP REBIUN DE PROYECTOS DIGITALES  
VII JORNADAS DE OS REPOSITORIOS  
XI COLOQUIO INTERNACIONAL DE CIENCIAS DE LA DOCUMENTACIÓN

2017



Infrastructures for Open Science in Europe: the power of Repositories

**Pedro Principe**  
Universidade do Minho

# Infrastructures for Open Science in Europe: the power of Repositories

Salamanca, 25 October 2017



Universidade do Minho



**ECA**

ECOSISTEMAS DEL  
CONOCIMIENTO  
ABIERTO



# REPOSITARIOS COMO INFRAESTRUCTURA

1

2

## OpenAIRE y EL PODER DE LOS REPOSITARIOS



**Nuestros repositorios  
pueden ser los tesoros de  
nuestra institución...**

**Pero no pueden ser  
islas del tesoro...**





Cada repositorio individualmente puede ser valioso para su institución o comunidad.

**Pero cada repositorio solo tiene un valor muy limitado...**

**EL VALOR DE LOS REPOSITARIOS**



**El valor real de los repositorios recae en el potencial de interconexión para crear una red de repositorios, una red que puede proveer un acceso unificado a los resultados de la investigación y son (re-) utilizados por las máquinas y los investigadores.**



# Contenidos OA

# Interoperabilidad



Para lograr este potencial de crear un cuerpo unificado de resultados de la investigación necesitamos...



# Logramos las redes... pero...

*Devolver el control del sistema de comunicación científica a la comunidad investigadora y académica!*

**Fortalecer y ampliar el papel de la institución en la comunicación científica.**

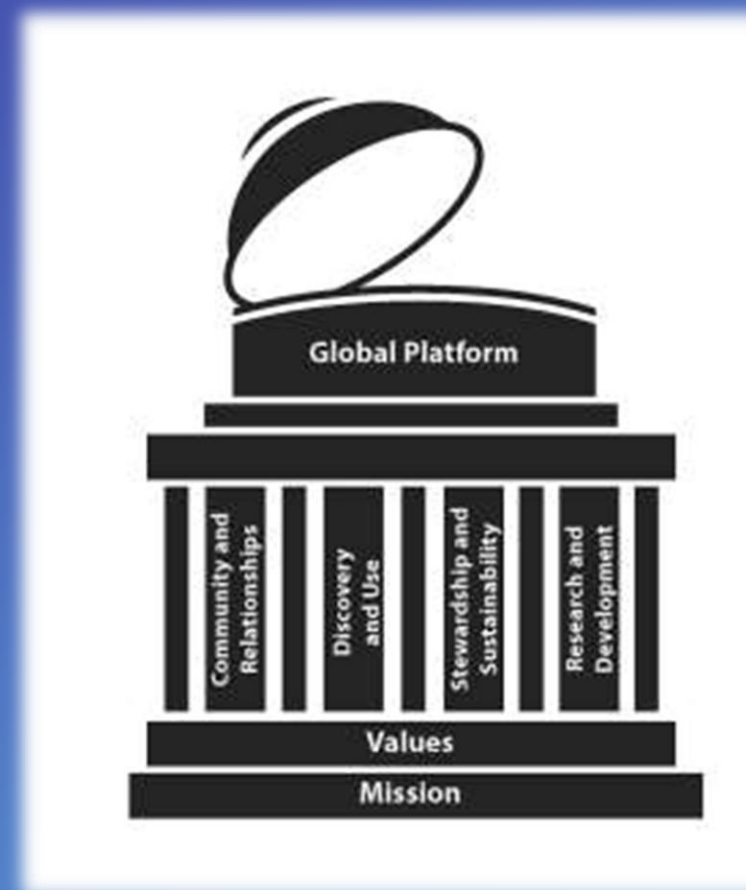
**Los repositorios son herramientas importantes para actualizar esta visión.**

“

“

“... The MIT Libraries must operate as an **open, trusted, durable, interdisciplinary, interoperable content platform** that provides a foundation for the entire life cycle of information for collaborative global research and education.”

”



”

*Futuro de las Bibliotecas, Oct 2016:  
Massachusetts Institute of Technology*

# Next generation repositories

## Visión

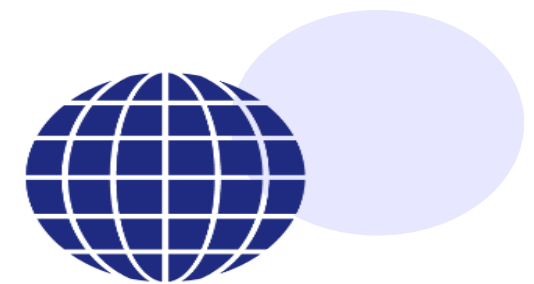
Posicionar los **repositorios** (y sus instituciones) como la **base** para una **infraestructura** de comunicación académica y científica **distribuida** y **globalmente interconectada**.



# Next generation repositories

## Dos ideas central para actualizar nuestra vision

- **Mejorar la funcionalidad de los repositorios:**
  - Ser de, y no solamente estar en, La Web.
  - Interoperabilidad global.
  - Repositorios pro-activos.
- **Suportar el desarrollo de servicios de valor agregado**  
*(recommender Systems, Open Peer Review, Usage stats).*



# Next generation repositories

content, metadata and people

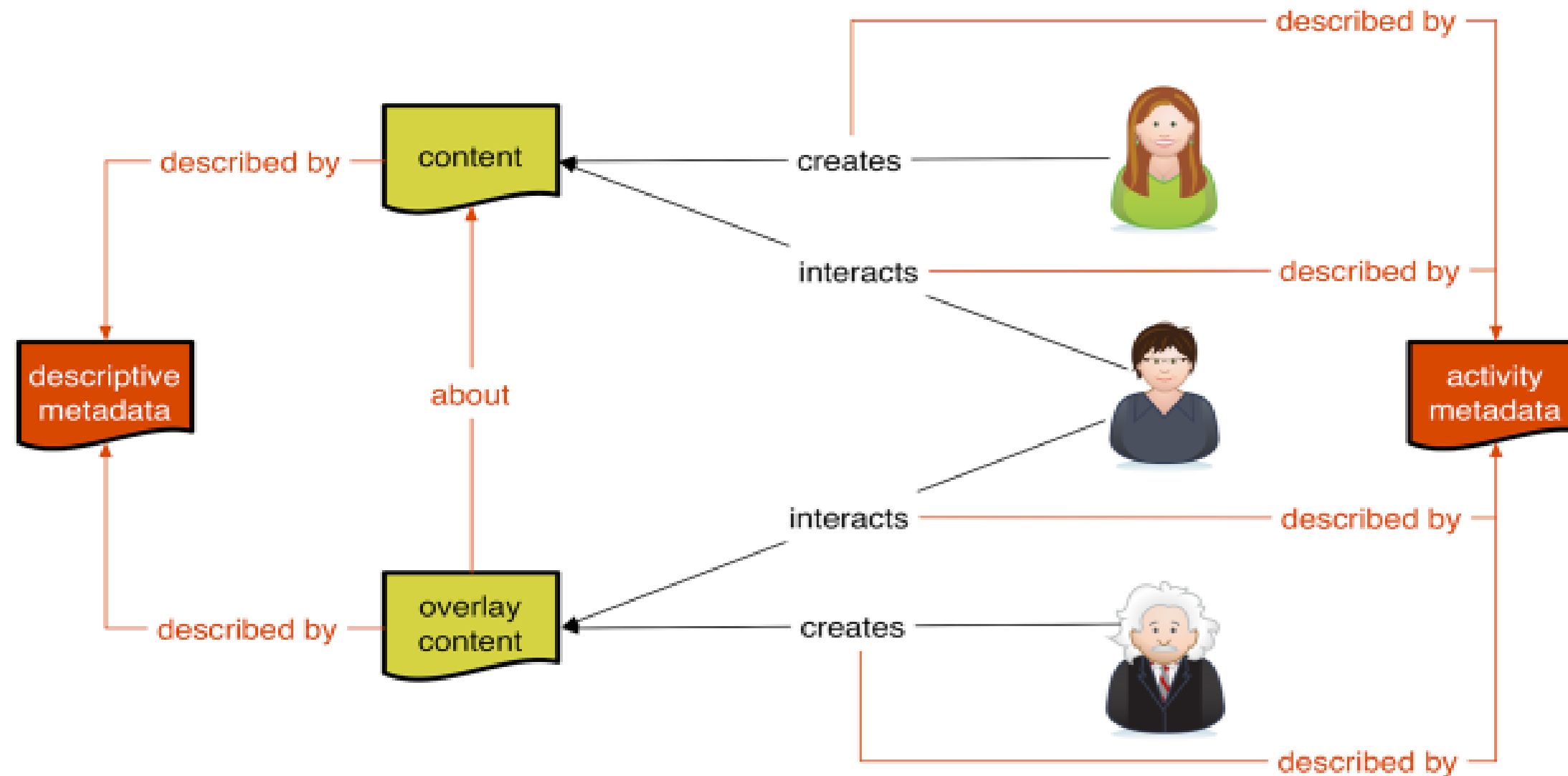
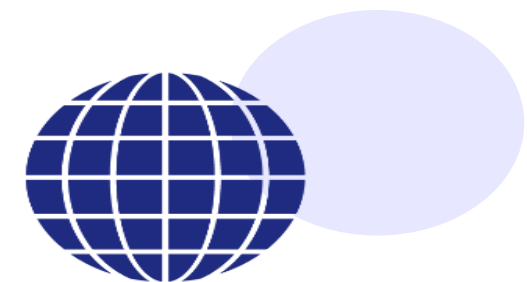


Diagram by Herbert Van de Sompel



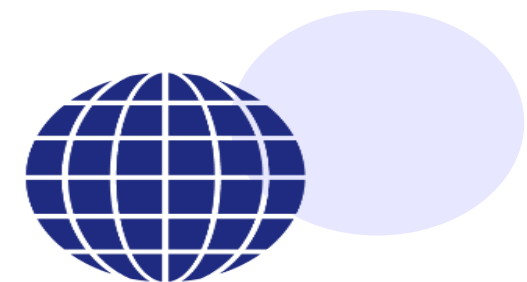
# Next generation repositories

repositories must be *deeply connected*

---

- **outgoing:**
  - individual *content resources*
    - directly accessible on the network
  - individual *metadata records*
    - not just in batches
  - individual *users*
    - as part of a variety of professional and social networks
- **incoming:**
  - using all appropriate *global identifier systems*
  - accepting *automated deposit* of content and data from other systems (e.g. *scientific instruments*)
  - allowing *external services* to interact with content
    - content mining
    - annotation services
    - etc.

Slide: Paul Walk






**It is potential for interoperation which elevates the institutional repository from a local system, to a networked system"**

*Repositories as Infrastructure* – Paul Walk





**Tecnología en evolución.  
Fuerte apoyo de la Comunidad.  
Controlo distribuido.**

**Repositories as Infrastructure!**

# REPOSITARIOS COMO INFRAESTRUCTURA

---

Públicos diversos.

---

Diferentes funciones.

---

Híbridos (catch-all).

---

Redes.



Universidade do Minho



**ECA**  
ECOSISTEMAS DEL  
CONOCIMIENTO  
ABIERTO



# OpenAIRE

Una mirada:

Services and recent/future developments  
for funders and content providers.



## one stop shop for OpenAIRE data providers

for friends...  
"the repository managers dashboard"

*Dashboard for content providers*

## Services for Funders

Align Open Science policies.  
Sync infrastructures.  
Monitor and report.

*OpenAIRE helps you measure  
research funding impact and  
supports monitoring of Open Access*

## Usage metrics for OpenAIRE Content Providers

A Usage Statistics Hub for  
Responsible Metrics

*Usage Statistics Service*



# OpenAIRE

## Open Science as-a-Service

OpenAIRE-connect

*Research Community Dashboard*

*Catch-All-Notification Broker*

## OpenAIRE Advancing Open Scholarship

OpenAIRE Advance 2018-2020

*OpenAIRE advance will build on the  
outcomes of the COAR Next  
Generation Repositories*

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



# one stop shop for OpenAIRE data providers

for friends...

“the repository managers dashboard”

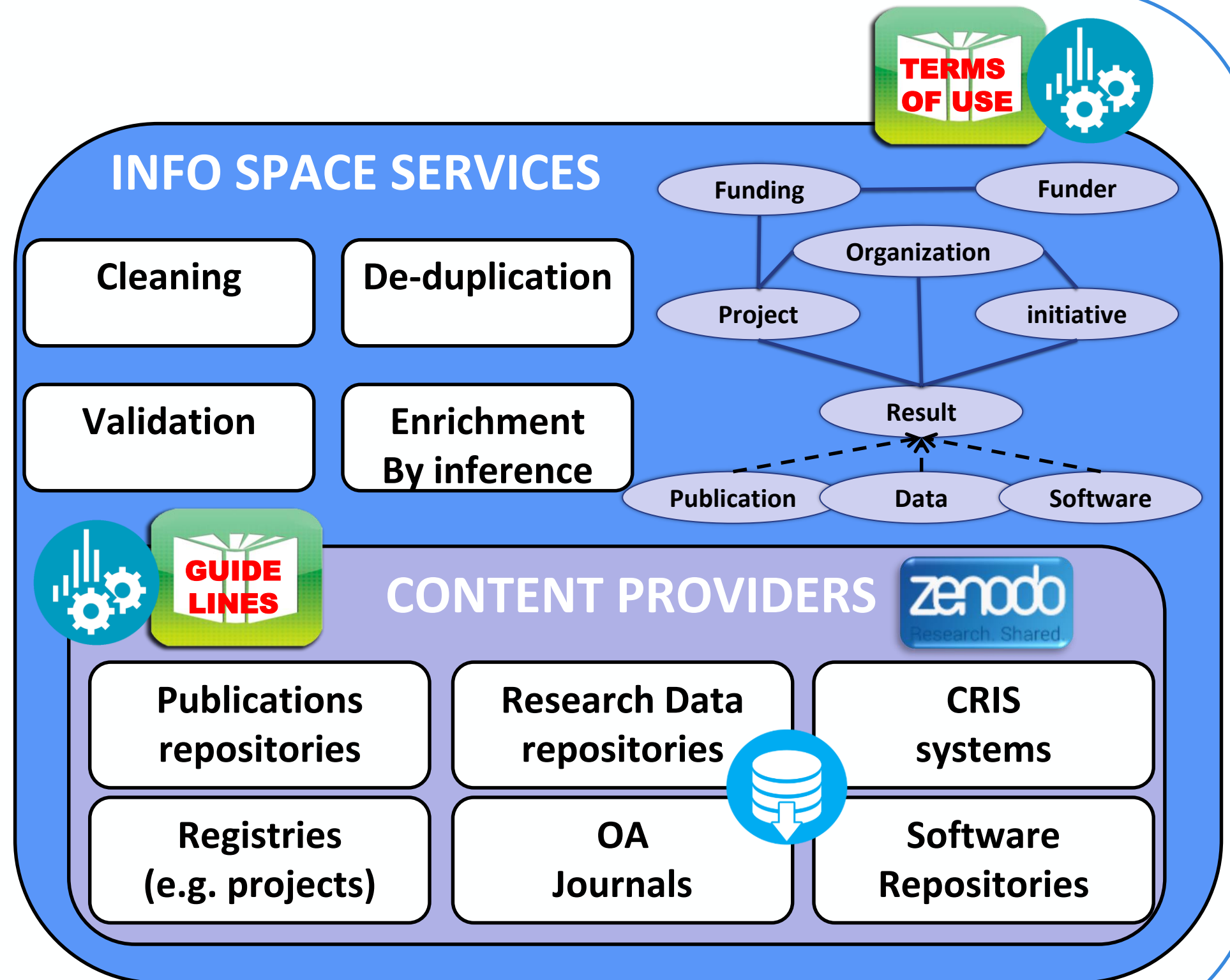
*Dashboard for content providers*

# OpenAIRE's e-infrastructure Commons

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.*



# Literature broker service

## THE CHALLENGE

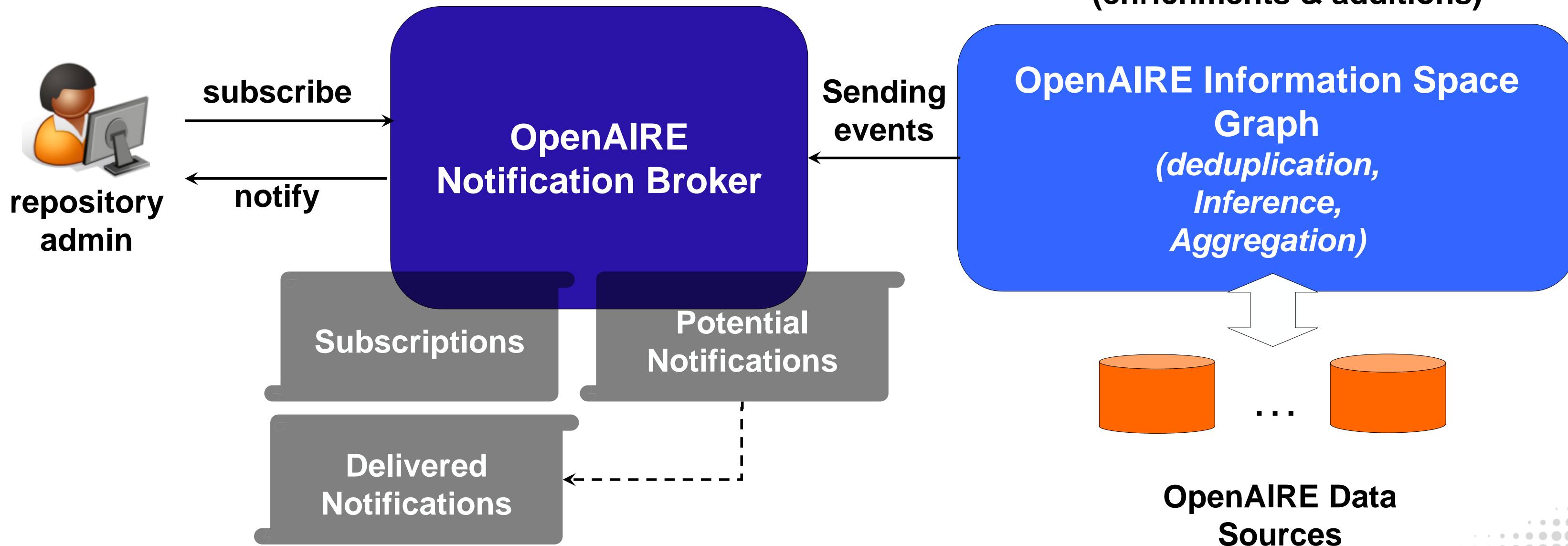
- Enrichment is straightforward
  - Harvesting from repository and return to repository its records if they have been “enriched” by deduplication and/or inference
- Addition is less obvious
  - Based on relationships, in turn identified by inference algorithms
  - Must be augmented with notion of “trust” to enable “tuning” options in order to reduce false positive notifications

# OpenAIRE Broker sketch

**Event (potential notification):**

- *Message*
- *Topic*
- *TargetRepository*
- *Trust*

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



# Subscription & notification

**Repositories can subscribe to the service and receive notifications about records of potential interest and specify**

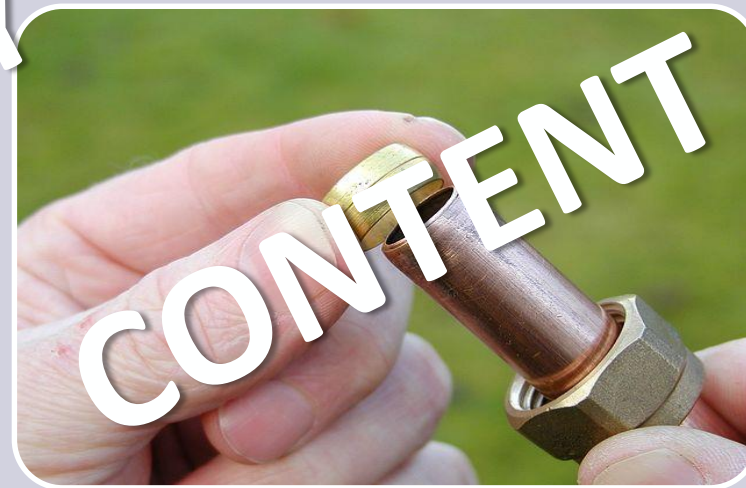
- what metadata fields they would like to be notified of
- how to be notified.

**The service can notify the repositories in different ways**

- via custom (OpenAIRE defined) repository APIs for metadata ingestion
- via email to the repository managers and via web interface.

**Broker services available via a specific dashboard  
for content providers...**

# Content Provider Dashboard



**REGISTER  
&  
UPDATE**

registration service

**VALIDATE  
&  
MONITOR**

validation service

**ENRICHMENTS  
&  
NOTIFICATIONS**

broker service

**METRICS  
USAGE  
STATISTICS**

usage stats service

# Dashboard Functionalities

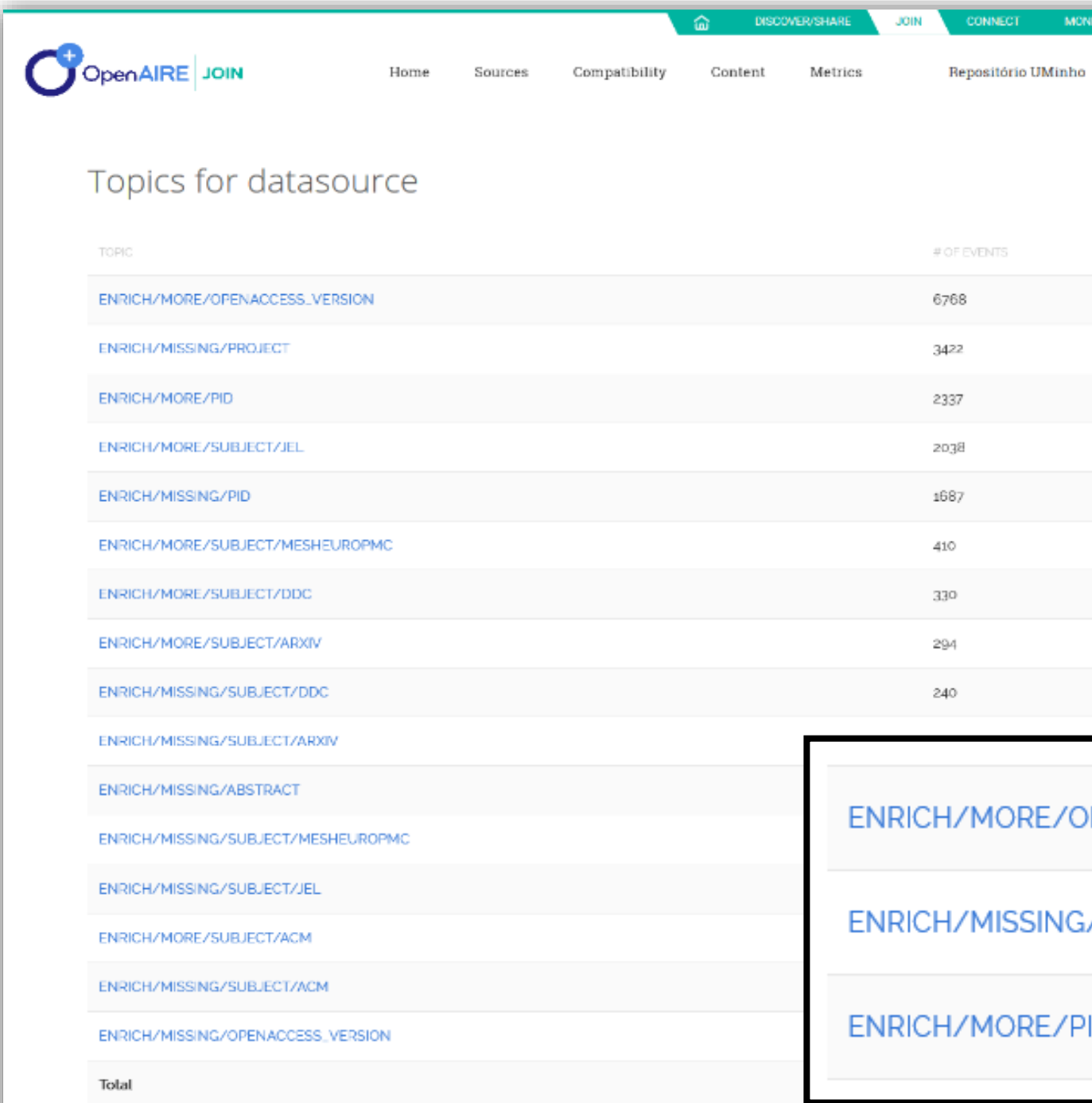
- ✓ Data source registration
- ✓ Validation (against guidelines)
- ✓ **Content enrichment and fixing**
- ✓ Content notifications
- ✓ Data source usage statistics

The screenshot shows the OpenAIRE dashboard with a navigation bar at the top containing 'DISCOVER/SHARE', 'JOIN', 'CONNECT', 'MONITOR', and 'DEVELOP'. Below this is a secondary navigation bar with 'Home', 'Sources', 'Compatibility', 'Content', 'Metrics', 'Repositório UMinho', and 'LOGOUT'. The main content area features four cards:

- Register:** Register data sources in the OpenAIRE infrastructure. (Image: person on a beach)
- Validate:** Validate data sources against OpenAIRE guidelines. (Image: landscape with a stone wall)
- Notifications:** View notifications to enrich the metadata and the content. (Image: mountain landscape)
- Metrics:** View aggregated, cleaned usage statistics for repository access. (Image: person on a rocky shore)

This screenshot shows the 'Enrich Your Content - Browse Events' page. It features the OpenAIRE logo and navigation links for 'Home', 'Sources', and 'Compatibility'. The main heading is 'Enrich Your Content - Browse Events', followed by the instruction 'CHOOSE THE DATASOURCE FOR WHICH YOU WOULD LIKE TO VIEW EVENTS'. A red button with a white starburst icon and the text 'RepositórioUM' is visible. To the right, it displays 'Universidade do Minho: RepositoriUM (18256 events)'.

# Content events - enrichments & alerts

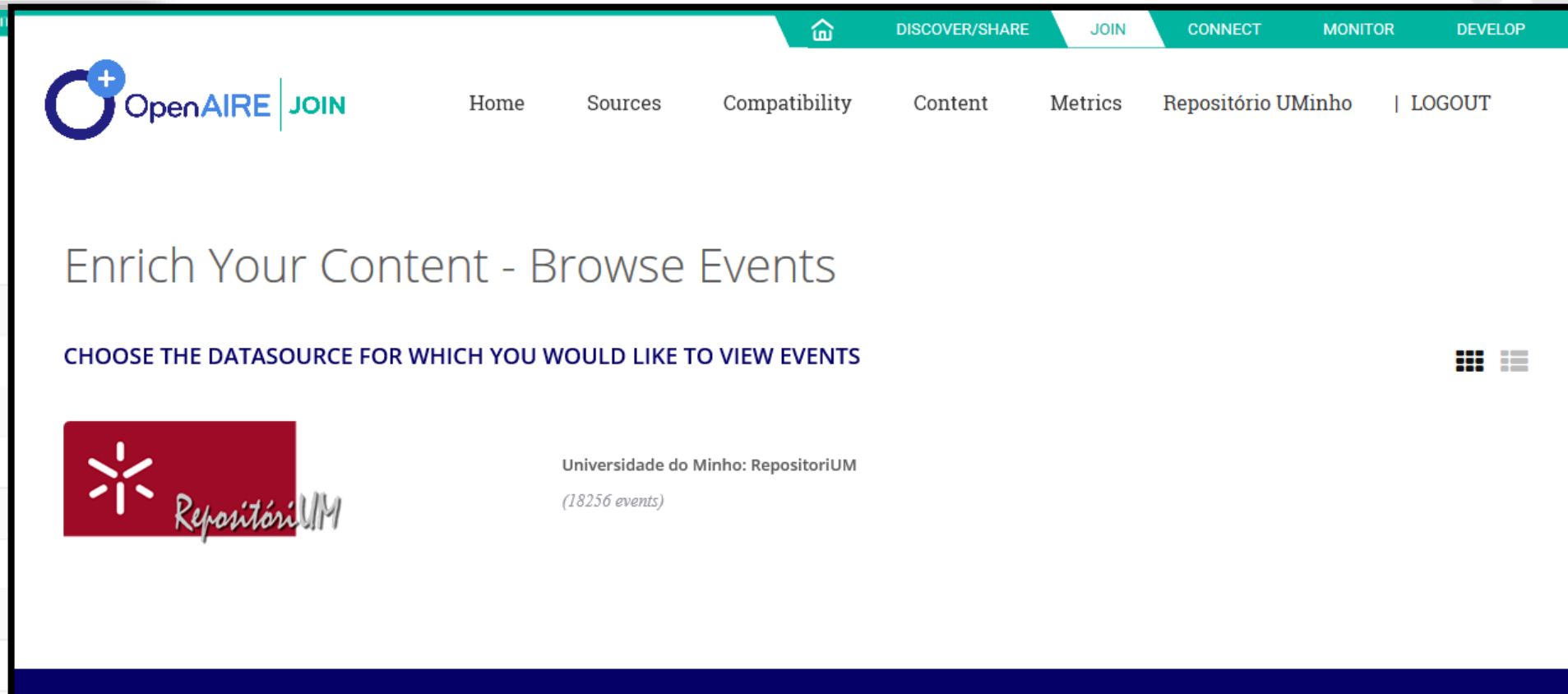


OpenAIRE JOIN

Home Sources Compatibility Content Metrics Repositório UMinho

### Topics for datasource

TOPIC	# OF EVENTS
ENRICH/MORE/OPENACCESS_VERSION	6768
ENRICH/MISSING/PROJECT	3422
ENRICH/MORE/PID	2337
ENRICH/MORE/SUBJECT/JEL	2038
ENRICH/MISSING/PID	1687
ENRICH/MORE/SUBJECT/MESHEUROPMC	410
ENRICH/MORE/SUBJECT/DDC	330
ENRICH/MORE/SUBJECT/ARXIV	294
ENRICH/MISSING/SUBJECT/DDC	240
ENRICH/MISSING/SUBJECT/ARXIV	
ENRICH/MISSING/ABSTRACT	
ENRICH/MISSING/SUBJECT/MESHEUROPMC	
ENRICH/MISSING/SUBJECT/JEL	
ENRICH/MORE/SUBJECT/ACM	
ENRICH/MISSING/SUBJECT/ACM	
ENRICH/MISSING/OPENACCESS_VERSION	
Total	




OpenAIRE JOIN

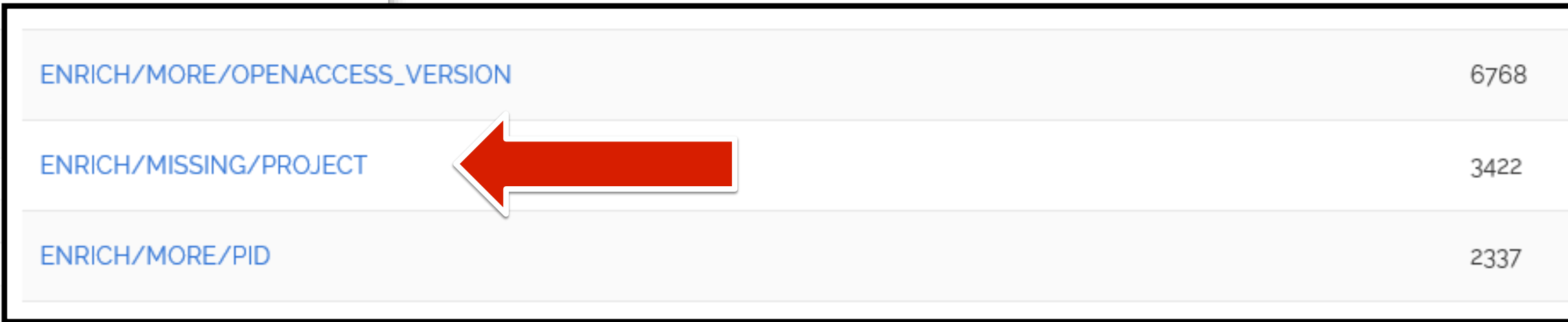
Home Sources Compatibility Content Metrics Repositório UMinho | LOGOUT

## Enrich Your Content - Browse Events

CHOOSE THE DATASOURCE FOR WHICH YOU WOULD LIKE TO VIEW EVENTS



Universidade do Minho: RepositoriUM  
(18256 events)



ENRICH/MORE/OPENACCESS_VERSION	6768
ENRICH/MISSING/PROJECT	3422
ENRICH/MORE/PID	2337

**DEMO**

# Science. Set free.

What we do

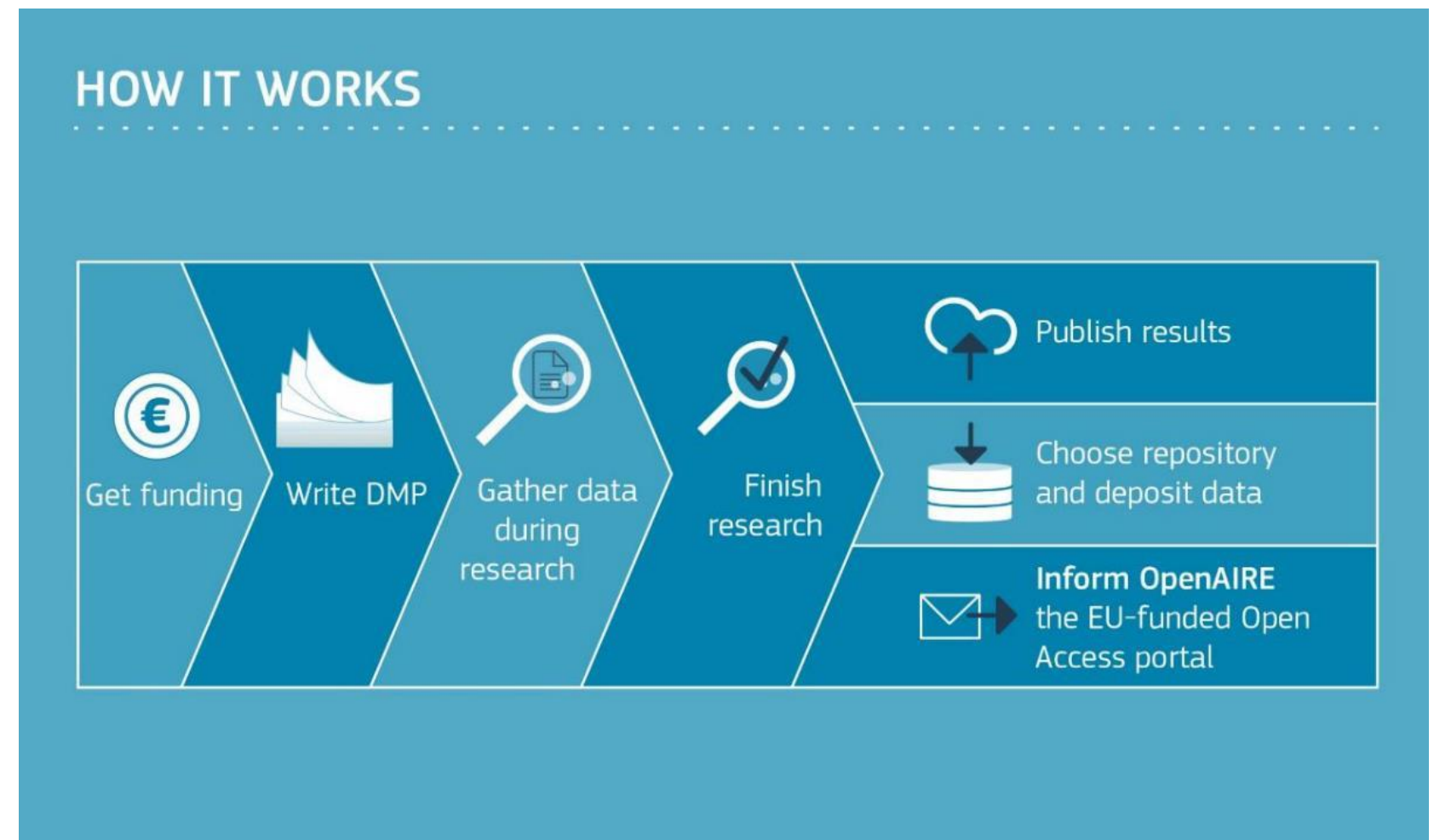
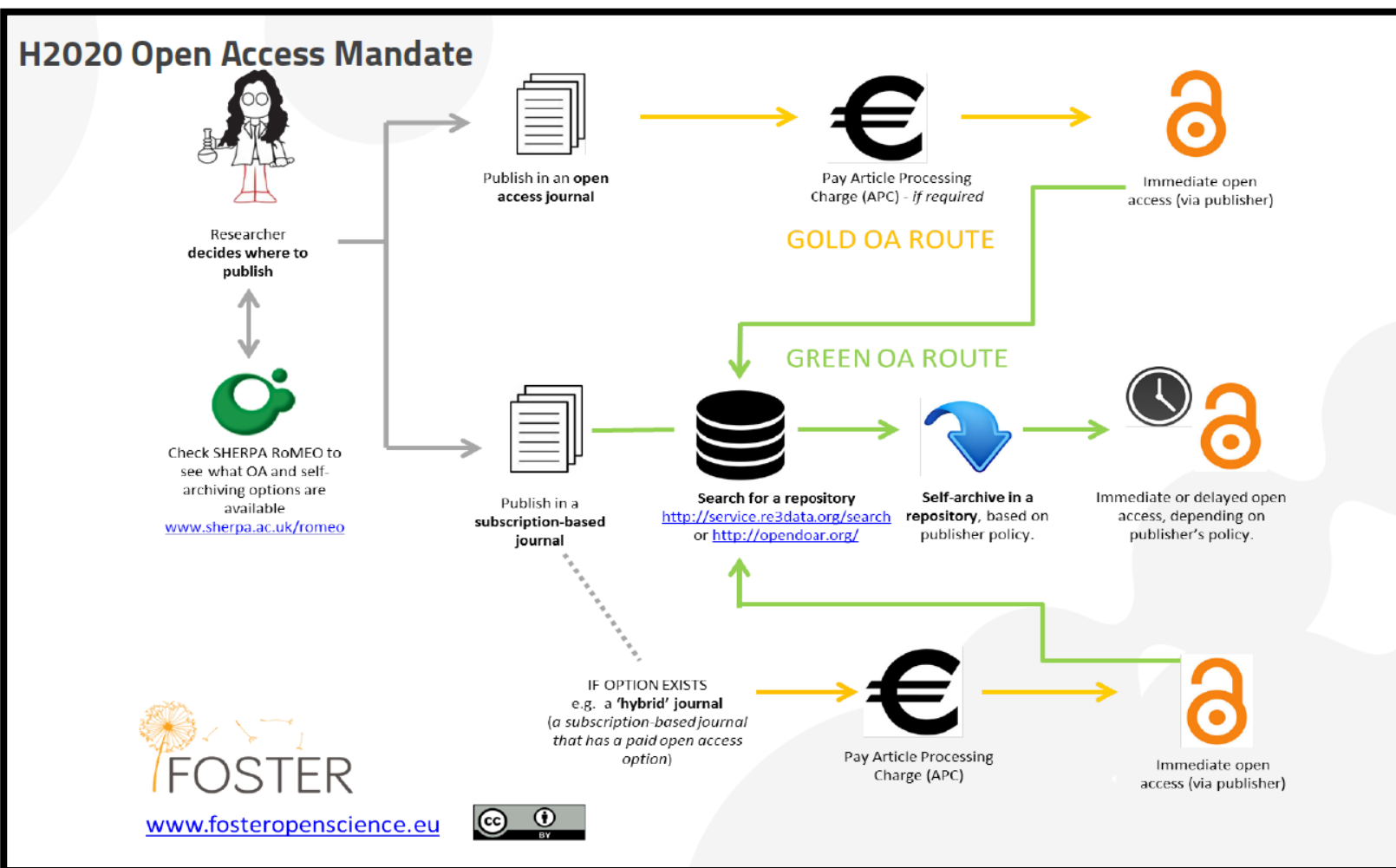
A lighthouse on a cliff overlooking the ocean under a cloudy sky. The lighthouse is white with a black top section and is illuminated. The ocean is dark with white foam from waves crashing against rocks in the foreground. The sky is overcast with grey clouds.

# Services for Funders

Align Open Science policies.  
Sync infrastructures.  
Monitor and report.

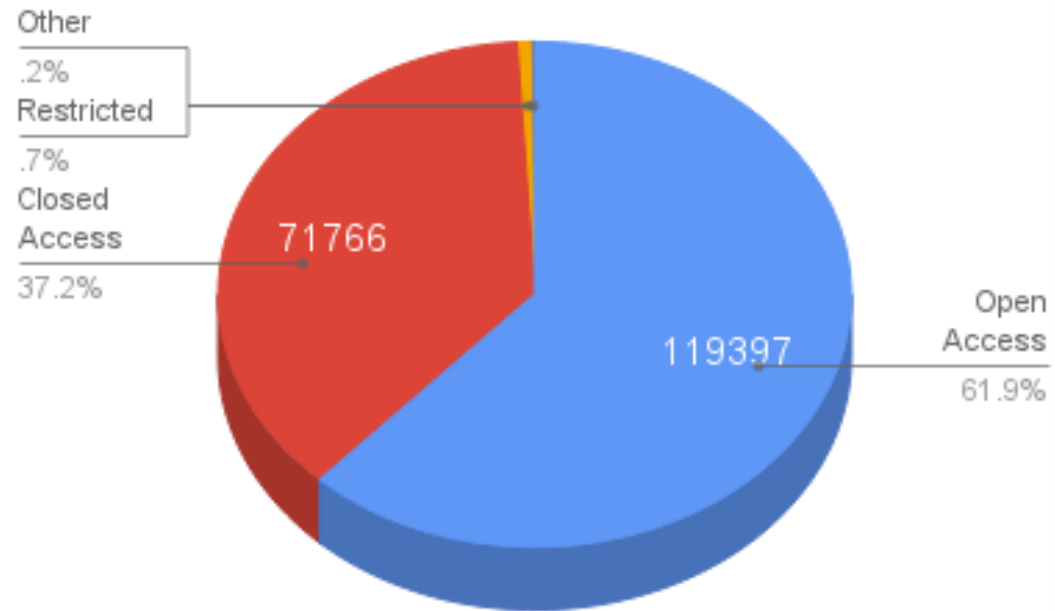
*OpenAIRE helps you measure  
research funding impact and  
supports monitoring of Open Access*

# Open Access & Open Research Data in H2020

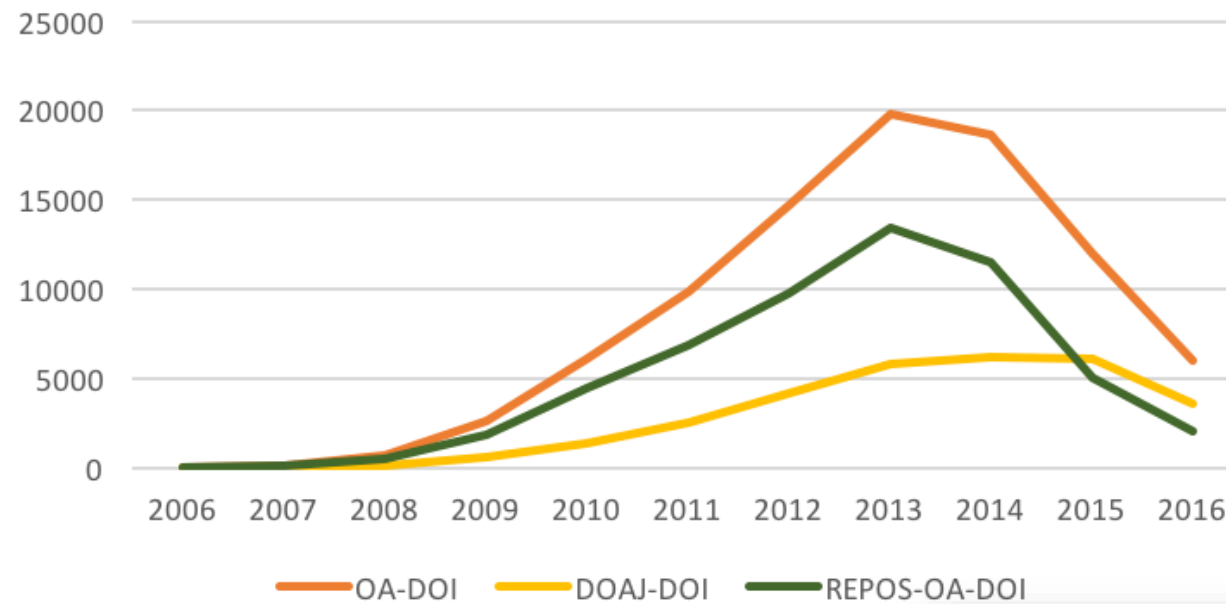


# OpenAIRE monitoring for EC

FP7 publications breakdown by access mode

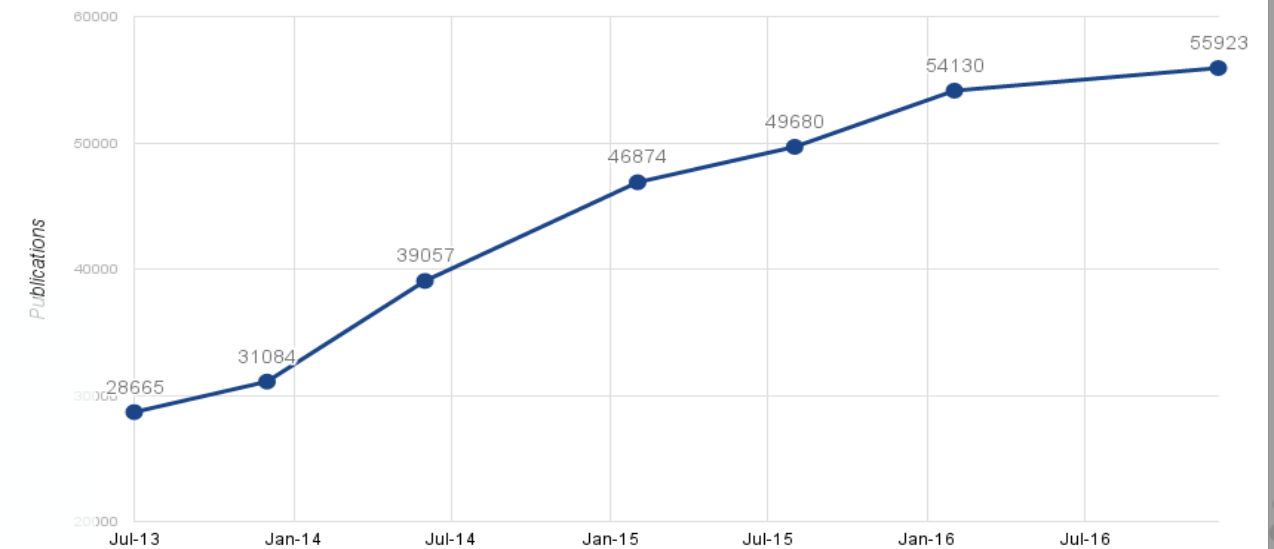


FP7 Gold - Green



OA with DOIs

FP7 publications in institutional repositories



# EC projects and funding streams pages and stats

**OpenAIRE** PARTICIPATE SEARCH

## H2020 Monitoring

19,181 publications in 2,951 projects (from a total of 14,260)  
8,408 are OA, 99 are restricted and 117 are still in embargo [Graph](#)

Publications in H2020 Projects through the years

Year	Publications
2014	~100
2015	~2,800
2016	~11,000
2017	~5,800

ABOUT PUBLICATIONS

- TOP 30 PROJECTS IN PUBLICATIONS
- H2020 PUBLICATIONS BY TYPE

ABOUT DATA PROVIDERS

- TOP 20 DATA PROVIDERS FOR H2020 PUBLICATIONS
- H2020 PUBLICATIONS BY DATA PROVIDER TYPE

## ERC statistics

78,227 publications in 3,921 projects (total of 4,561 projects)  
53,949 are OA, 382 are restricted and 89 are still in embargo [Graph](#)

ERC publications through the years

Year	Publications
2007	~100
2008	~200
2009	~500
2010	~1,500
2011	~3,500
2012	~13,000
2013	~20,000
2014	17,592
2015	~18,000
2016	~8,000
2017	~8,000

ABOUT PUBLICATIONS

TOP 30 PROJECTS IN PUBLICATIONS

ERC Publications by project (top 30)

Project	Publications
ATMICLE	~300
PROSPERITY	~260
SUPERFIELD...	~250
MUSEUM	~250
CM	~250
LASCANIX	~250
COUNTACTING	~240
SMCCOM	~230
DYNAMO	~220
TEMPURVIDE...	~210
ALPAM	~200
WATERFIELD...	~190
AMIMOS	184
MOJANZSP...	~180
COBRY	~170
IMBALANCE...	~160
SAW	~160
ACTIDISK	~160
NUMERNAVE...	~160
ARCTIC DOM...	~160
FLAVOUR	~160
CFSEE	~160

## FP7 statistics

220,828 publications in 15,712 projects (from a total of 25,870)  
141,829 are OA, 3,476 are restricted and 344 are still in embargo [Graph](#)

For SC39:  
31,889 publications in 1,498 projects (total of 1,907)  
22,840 are OA [Graph](#)

FP7 TIMELINE

To get more detailed views please make your selections from either Programme or Scientific Area:

Programme

- SP1-Cooperation
- SP2-Ideas
- SP3-People
- SP4-Capacities
- SP5-Euratom

Scientific Area

- COH
- ENERGY
- ENV
- ERC
- Fission
- Fusion
- GA

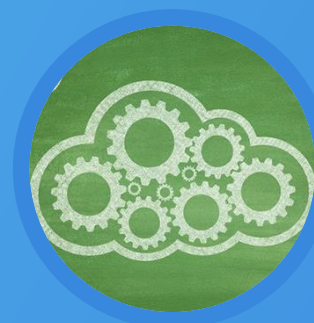
FP7 publications through the years

Year	Publications
2007	~100
2008	~2,000
2009	~6,000
2010	~15,000
2011	~25,000
2012	~35,000
2013	~47,000
2014	~36,000
2015	~19,000
2016	~14,000
2017	~1,000

# Funders related activities in OpenAIRE



**Align and support OA policies**



**APIs with projects list for repository software**



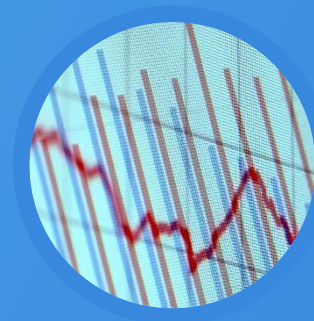
**Sync infrastructures and Support national e-infrastructures**



**Text mining and inference**



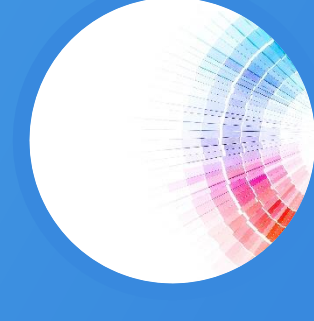
**Guidelines for metadata funding info**



**Statistics and reporting**



**Monitor mandate compliance**



**Analytics and trends**

# BETA

# www.openaire.eu

The screenshot shows the OpenAIRE website interface. At the top left is the OpenAIRE logo. Below it is a navigation bar with tabs for 'Publications', 'Research Data', 'Projects', and 'People'. A search bar is visible with the text 'Search keywords'. A dropdown menu titled 'FUNDER' is open, displaying a list of funding organizations and their associated counts. The list includes:

- National Institutes o... (1682692)
- National Science Foun... (497646)
- Research Council UK (70193)
- Swiss National Scienc... (65490)
- European Commission (40964)
- Fundação para a Ciênc... (37277)
- National Health and M... (25415)
- Australian Research C... (24269)
- Netherlands Organisat... (24180)
- Suomen Akatemia (23932)
- Deutsche Forschungsge... (21796)
- Türkiye Bilimsel ve T... (16609)
- Austrian Science Fund... (13642)
- Wellcome Trust (12753)
- Comisión Nacional de ... (8613)
- Gobierno de España (6674)
- Science Foundation Ir... (3411)
- Ministry of Science, ... (2120)
- Croatian Science Foun... (881)
- Ministry of Education... (777)
- Tara Expeditions Foun... (5)

The screenshot shows the OpenAIRE website interface, similar to the first one. The 'FUNDER' dropdown menu is open, displaying a different list of funding organizations and their associated counts. The list includes:

- National Institutes o... (1542722)
- National Science Foun... (497646)
- Swiss National Scienc... (66879)
- European Commission (41053)
- Fundação para a Ciênc... (37277)
- National Health and M... (24354)
- Australian Research C... (24268)
- Netherlands Organisat... (24180)
- Austrian Science Fund... (13619)
- Wellcome Trust (12753)
- Science Foundation Ir... (3411)
- Ministry of Science, ... (2120)
- Croatian Science Foun... (881)
- Ministry of Education... (777)

Funder	Country	Status
European Commission (FP7 & H2020)	EU	Production
Portuguese Foundation for Science and Technology (FCT)	Portugal	Production
Wellcome trust (WT)	UK	Production
National Science Foundation (NSF)	US	Production
National Health and Medical Research Council (NHMRC)	Australia	Production
Australian Research Council (ARC)	Australia	Production
Ministry of Education, Science and Technological Development (MESTD)	Serbia	Production
Ministry of Science Education and Sport (MZOS/MSES)	Croatia	Production
Croatia Science Foundation (HRZZ/CSF)	Croatia	Production
Netherlands Organisation for Scientific Research (NWO)	Nederlands	Production
Science Foundation Ireland (SFI)	Ireland	Production
National Institute of Health (NIH)	US	Production
Swiss National Science Foundation (SNSF)	Switzerland	Production
Austrian Science Fund (FWF)	Austria	Production
Research Council UK (RCUK)	UK	BETA
Tara Expedition Foundation (TARA)	France	BETA
Scientific and Technological Research Council of Turkey (Tubitak)	Turkey	<b>BETA</b>
Academy of Finland	Finland	BETA
CONICYT	Chile	BETA
<b>Ministry of Economy, Industry and Competitiveness</b>	<b>Spain</b>	<b>BETA</b>
German Research Foundation (DFG)	Germany	Pilot

# Funder integration in OpenAIRE: simplified workflow



1. Obtain list of projects
2. Data analysis for integration and inference module
3. Integration on DB
4. IIS module integration

5. Inference
6. Data provision on beta

7. Quality check  
Results available on the OpenAIRE portal

# ESPAÑA: 2 funding streams, 4623 projects = 4622 publications (provisional results – testing)

The screenshot shows the OpenAIRE search results page for the filter 'Funder: Gobierno de España'. The page displays 4622 documents on page 1 of 463. The search filters section on the left lists various criteria such as 'Also funded by', 'Funding Stream', 'Project', 'Publication Year', 'Access Mode', 'Document Type', 'Document Language', and 'Data Provider'. The main content area shows two search results:

- Secondhand tobacco smoke exposure in open and semi-open settings: a systematic review** (2013)  
Projects: SGOV | Cáncer (RD12/0036/0053)  
Background: Some countries have recently extended smoke-free policies to particular outdoor settings; however, there is controversy regarding whether this is scientifically and ethically justifiable. Objectives: The objective of the present study was to review research on secondhand smoke (SHS) exposure in outdoor settings. Data sources: We conducted different searches in PubMed for the period prior to September 2012. We checked the references of the identified papers, and conducted a similar...
- Optimization of Dengue Immunoassay by Label-Free Interferometric Optical Detection Method** (2014)  
Projects: SGOV | BIOCHIP KITS BASADOS EN CELDAS BIOFOTONICAS Y PLATAFORMAS AVANZADAS DE INTERROGACION OPTICA (TEC2012-31145)  
In this communication we report a direct immunoassay for detecting dengue virus by means of a label-free interferometric optical detection method. We also demonstrate how we can optimize this sensing response by adding a blocking step able to significantly enhance the optical sensing response. The blocking reagent used for this optimization is a dry milk diluted in phosphate buffered saline. The recognition curve of dengue virus over the proposed surface sensor demonstrates the capacity of th...

At the bottom of the page, there are two additional search results:

- Galic acid (3,4,5-trihydroxybenzoic acid, GA)** is a plant secondary metabolite, which shows antioxidant activity and is commonly found in many plant-based foods and beverages. Recent evidence suggests that oxidative stress contributes to the development of many human chronic diseases, including cardiovascular and neurodegenerative pathologies, metabolic syndrome, type 2 diabetes and cancer. GA and its derivative, methyl-3-O-methyl gallate (M3OMG), possess physiological and pharmacological act...
- Amyloids in solid-state nuclear magnetic resonance: potential causes of the usually low resolution** (2015)  
Projects: SGOV | NANOPARTICULAS MAGNETICAS PARA APLICACIONES TERANOSTICAS: MAGNETOLIPOSOMAS (MAT2012-36270-C04-03)  
Alba Espargaró; María Antón; Graciana Busquets, Joan Estelrich, Raimon Sabate Department of Physical Chemistry, School of Pharmacy, Institute of Nanoscience and Nanotechnology (IN2UB), University of Barcelona, Barcelona, Spain Abstract: Amyloids are non-crystalline and insoluble, which imply that the classical structural biology tools, i.e. X-ray crystallography and solution nuclear magnetic resonance (NMR), are not suitable for their analysis. In the last years, solid-state NMR...

# Funded projects metadata info in OpenAIRE



Collect **metadata** including project grantID from OpenAIRE compliant repositories



Metadata publications record enrichments by OpenAIRE **deduplication**



Link Publications to projects by **inference** (text mining procedures)



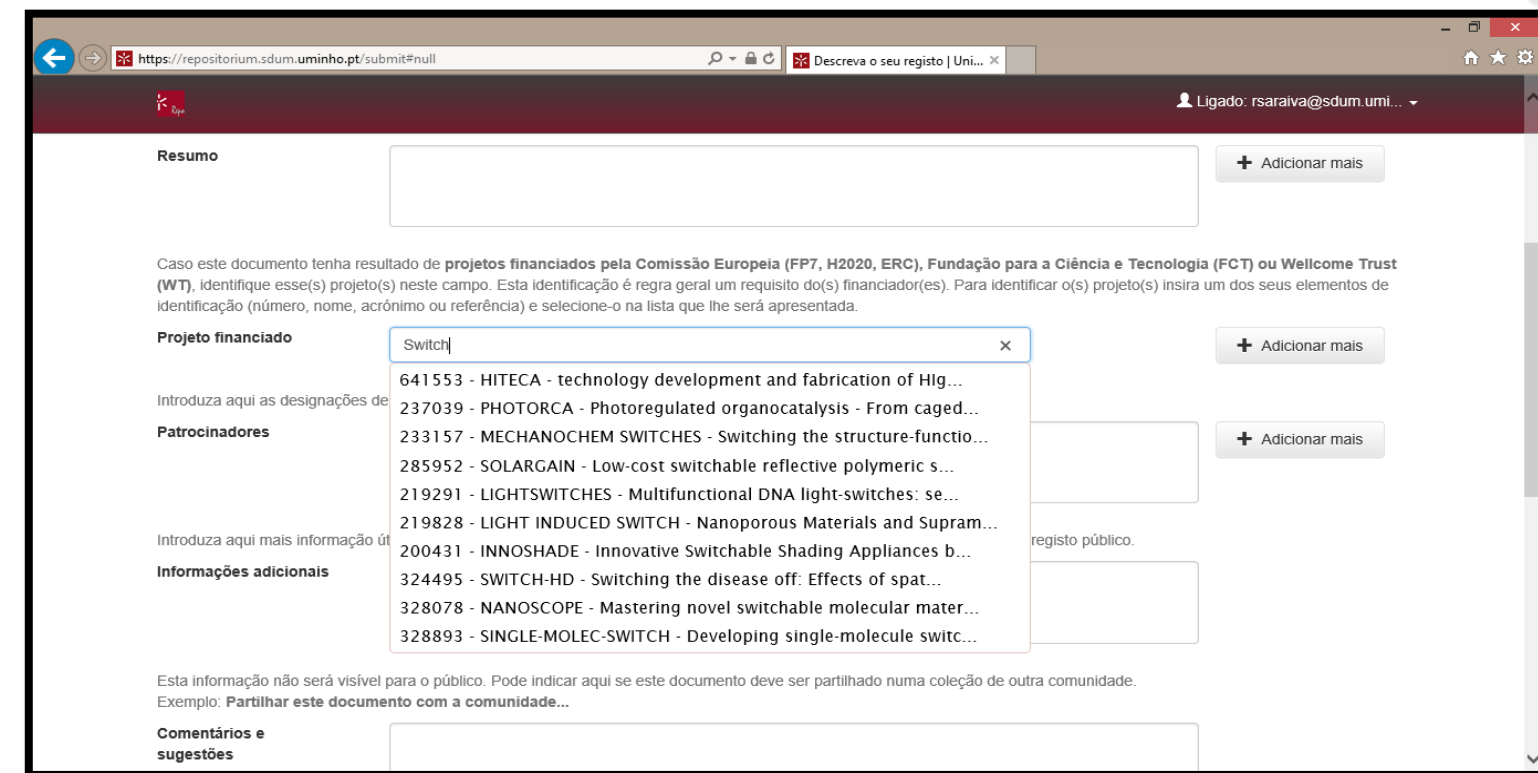
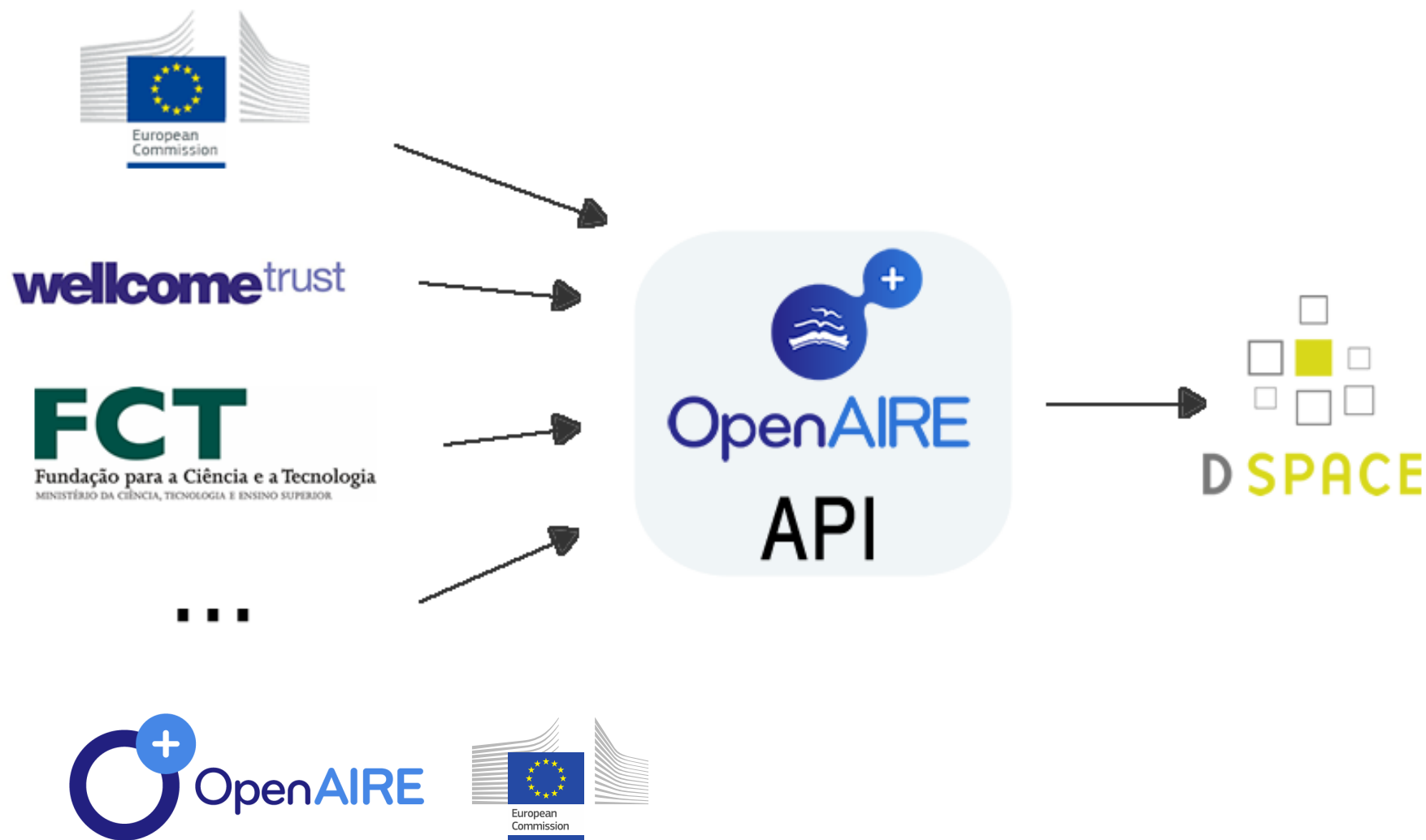
Link Publications to projects using the end-user service: **claim publications**



# OpenAIRE API (1/3)

The APIs offer custom access to metadata about projects funded by all funders integrated in the OpenAIRE info space (Dspace, Eprints, other repository platforms).

(E.g. Funders Projects List Dspace Add-on)



# Link research results portal facilities (2/3)

‘Claim’ functionality available for all national funders

OpenAIRE

HOME PARTICIPATE SEARCH MONITOR SUPPORT OPEN ACCESS

1 IDENTIFY PROJECT | 2 SELECT PUBLICATIONS/DATASETS | 3 SET ACCESS RIGHTS

### SELECT FUNDING AGENCY

1. Select funding agencies

Funder: [dropdown]

- Funder:
- Australian Research Council (ARC)
- Croatian Science Foundation (CSF)**
- European Commission
- Fundação para a Ciência e a Tecnologia, I.P.
- Ministry of Education, Science and Technological Development of Republic of Serbia
- Ministry of Science, Education and Sports of the Republic of Croatia (MSES)
- National Health and Medical Research Council (NHMRC)
- National Institutes of Health
- National Science Foundation
- Netherlands Organisation for Scientific Research (NWO)
- Science Foundation Ireland
- Wellcome Trust

### SELECT CONTEXTS(S)

1. Select community and category

Community: [dropdown] Category: [dropdown]

2. Select Contexts(s) [BROWSE](#)

Please select concept...

PROCEED

### Participate

- Deposit Publications & Data
- [Link Research Results](#)
- Validate / Register Repository
- Content policy

ABOUT PARTNERS NOADS

NEW FAC DO

Link publication or datasets to projects:  
Identify the project, select publications or  
datasets and set the access rights.

# Zenodo – extended grant support (3/3)

zenodo | Blog

About Blog Help Developers

## Extended grant support, ORCID and language field.

by Krzysztof Nowak on October 10, 2017

Today, we are introducing three additions to Zenodo:

1. Extended grant support
2. Language field
3. ORCID for authors

### Extended grant support (powered by OpenAIRE)

We are expanding our grants database with over 620,000 grants from 8 new funders such as the National Science Foundation (US) and Wellcome Trust (UK) - all thanks to [OpenAIRE's](#) ever growing grants database.

**Grants** European Commission (EU) **OpenAIRE-Connect** 731011 OpenAIRE - CONNECTing scientific results in s

National Science Foundation (US)

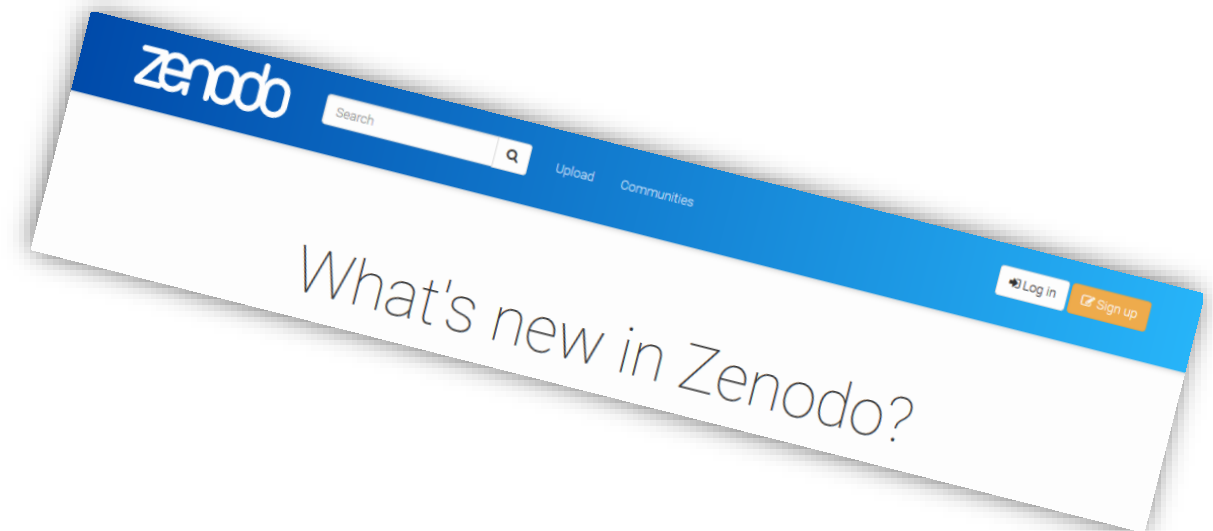
Optional. OpenAIRE-supported projects only. For other funding agencies, please use the *Additional Notes* field. Note: a human Zenodo curator will need to validate your uploads.

**1057396** **Open Access:** Conservation, Digitization and Interoperability of the Historic Non-Vertebrate Collections of the Texas Natural Science Center

[+ Add another grant](#)

So far we have only been supporting grants from the European Commission (FP7 and Horizon 2020). Today our dataset has grown and contains

- National Science Foundation (USA) - 497646 grants
- European Commission (EU) - 39409 grants
- Foundation for Science and Technology (Portugal) - 37277 grants



Funding recommended

Zenodo is integrated into reporting lines for research funded by the European Commission via [OpenAIRE](#). Specify grants which have funded your research, and we will let your funding agency know!

**Grants**

- Australian Research Council (AU)
- European Commission (EU)
- Fundação para a Ciência e a Tecnologia (PT)
- Ministarstvo Prosvete, Nauke i Tehnološkog Razvoja (RS) recommended
- Ministarstvo Znanosti, Obrazovanja i Sporta (HR) optional
- National Health and Medical Research Council (AU) optional
- National Science Foundation (US) optional
- Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NL) optional
- Wellcome Trust (UK) optional

Related/alternate identifiers recommended

Contributors optional

References optional

Journal optional

Conference optional

Book/Report/Chapter optional

funding acknowledgements, please use the *Additional Notes* field. Note: a human Zenodo curator will need to validate your uploads - you may experience a delay before it is available in OpenAIRE.





# Usage metrics for OpenAIRE Content Providers

A Usage Statistics Hub for  
Responsible Metrics

*Usage Statistics Service*

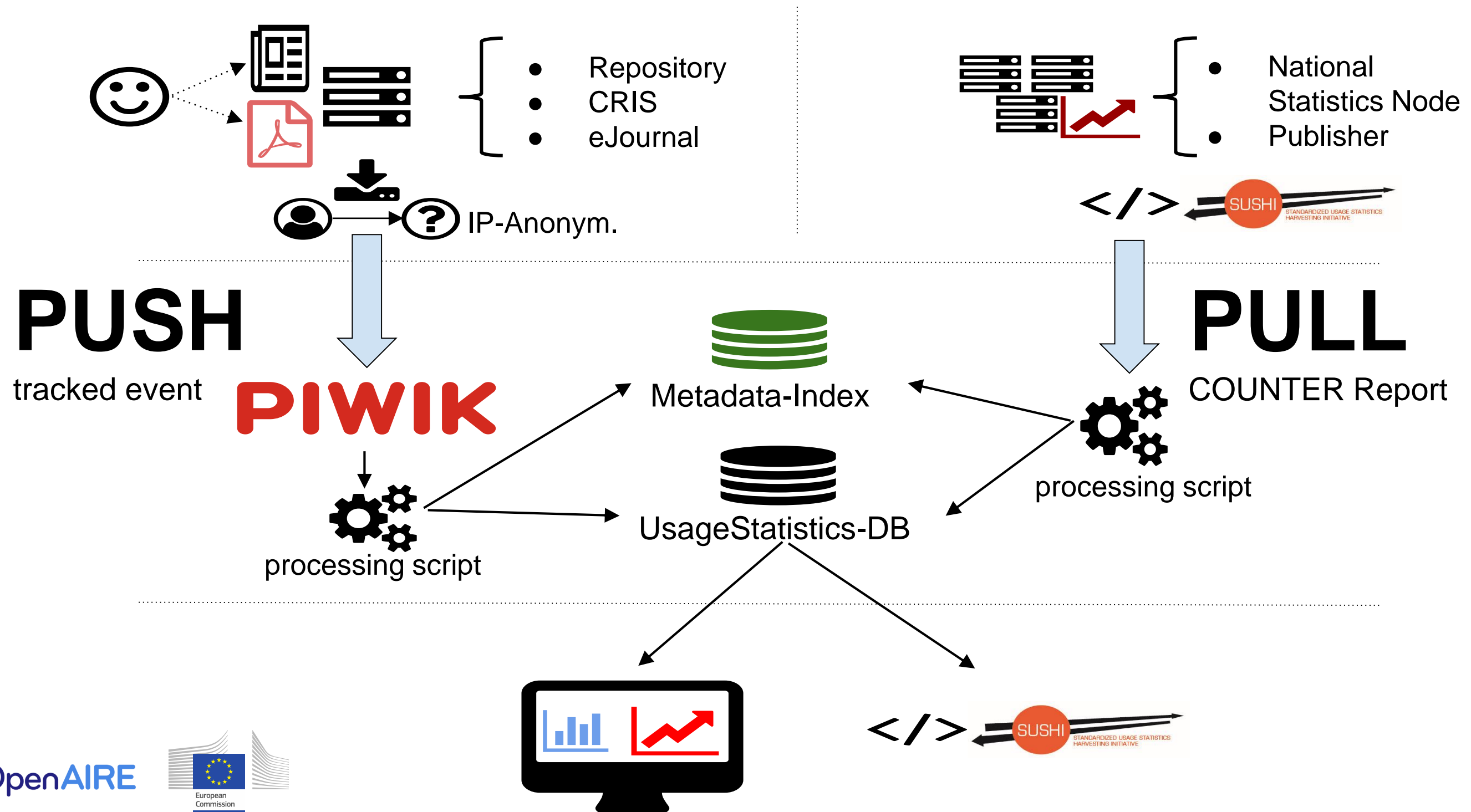
# Usage Statistics Service Features

- Tracking of views and downloads / collecting COUNTER reports
  - *Push or Pull* collection workflows.
- Anonymisation of IP-addresses.
- metadata deduplication enables accumulation of views and downloads for same documents
- COUNTER Code of Practice compatibility.
  - standards based usage statistics.
  - enables comparability with statistics from other data sources.

# Piwik Analytics platform

- World's leading open-source analytics platform.
- Valuable insights into website traffic and visitors activity.
- Piwik collects and stores *PII* (personally identifiable information).
- Keeps full data ownership and can control who has access.
- Robot filtering plugin.
- Compliant with EU regulations.
- Recommended by privacy organizations such as ULD (Germany) and CNIL (France).

# 2-Tiers Collection Workflows for Usage Statistics



# Tier-1: Push Usage Statistics Tracking Workflow

- An institutional repository is registered in Piwik.
- Two tracking options:
  - Client side: A small Javascript snippet embedded in web pages.
  - Server side: A plugin, using Piwik's HTTP API, developed for DSpace (patch) and Eprints.
- Usage Activity is tracked and logged at Piwik platform in real time.
- Information is transferred offline, using Piwik's API, to OpenAire's DBs for statistical analysis.
- Statistics are deployed via OpenAire's Portal or Sushi-Lite API.

# Tier-1 Usage Statistics Tracking Workflow

- Client Side Javascript snippet
  - Easy configurable.
  - Can be applied to any type of repository infrastructure.
  - Information can be blocked by javascript blockers.
  - “Direct” downloads are missed.
- Server Side plugin
  - More accurate tracking.
  - Tackles missing “direct” downloads problem.
  - A custom plugin should be developed for each repository infrastructure.

# Cleaning and Consolidation

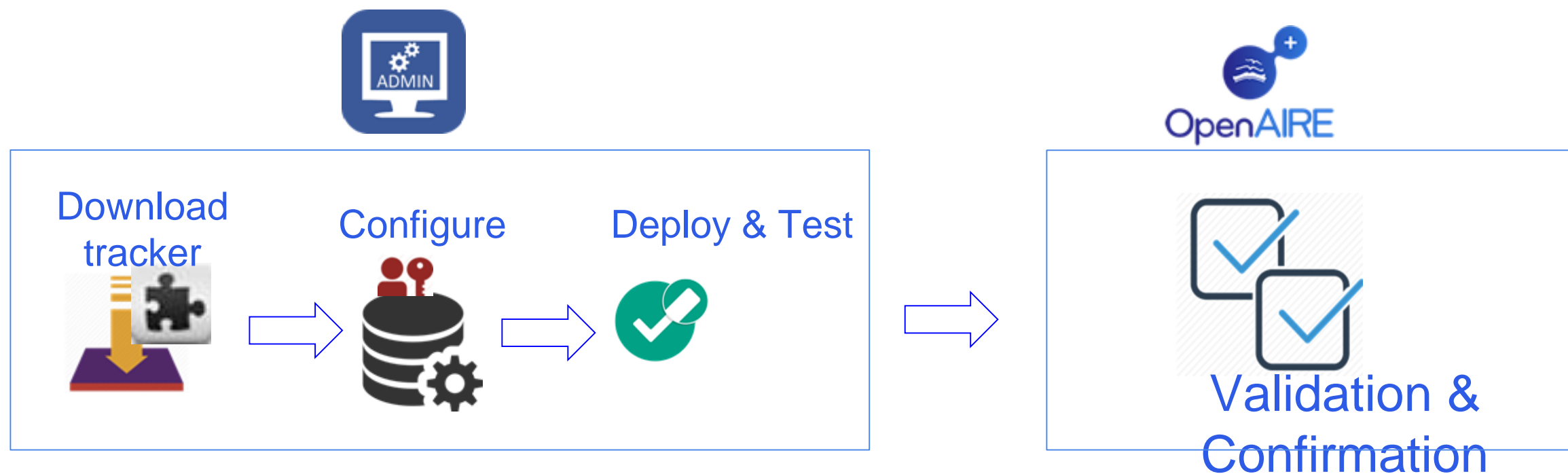
- Applying data processing rules according to COUNTER Code of Practice:
  - [ie. counting requests depending on session duration, tracing double-clicks](#)
- Bot filtering
  - [Piwik Bot Plugin](#)
  - [COUNTER Robots Working Group](#)
- Link of usage event with metadata record in OpenAIRE
- Accumulate views and counts of deduplicated records

## Tier-2: Collecting (Pull) Consolidated Usage Statistics Reports

- **Gathering of consolidated statistics reports from aggregation services, such as IRUS-UK, using protocols such as SUSHI-Lite.**
- **Statistics are stored to OpenAire's DB for statistical analysis.**
- **Statistics are deployed via OpenAire's Portal or Sushi-Lite API.**

# Metrics in the OpenAire Repository Manager Dashboard

- three steps to join OpenAIRE Usage Statistics
  - enable “usage metrics” for your data source
  - download & configure tracking plugin in your data source
  - confirmation by OpenAIRE once usage events are tracked in PIWIK
- or enter SUSHI endpoint to let OpenAIRE collect COUNTER reports



# Content Provider Dashboard - Start Page

The screenshot shows the OpenAIRE Content Provider Dashboard. At the top, there is a navigation bar with a home icon, 'DISCOVER/SHARE', 'JOIN', 'CONNECT', 'MONITOR', and 'DEVELOP'. Below this is a secondary navigation bar with 'Home', 'Sources', 'Compatibility', 'Content', and 'Metrics'. The user 'Jochen Schirrwagen' is logged in, with a 'LOGOUT' link. The main content area features four cards: 'Register' (with a folder icon), 'Validate' (with a downward arrow icon), 'Notifications' (with a network graph icon), and 'Metrics' (with a blue ribbon icon). The 'Notifications' card is highlighted with a white border.

**OpenAIRE | JOIN**

Home Sources Compatibility Content Metrics

Jochen Schirrwagen | LOGOUT

**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

# Enable Metrics for selected Datasource

OpenAIRE | JOIN

Home Sources Compatibility Content Metrics Jochen Schirrwagen | LOGOUT

DISCOVER/SHARE JOIN CONNECT MONITOR DEVELOP

## Metrics

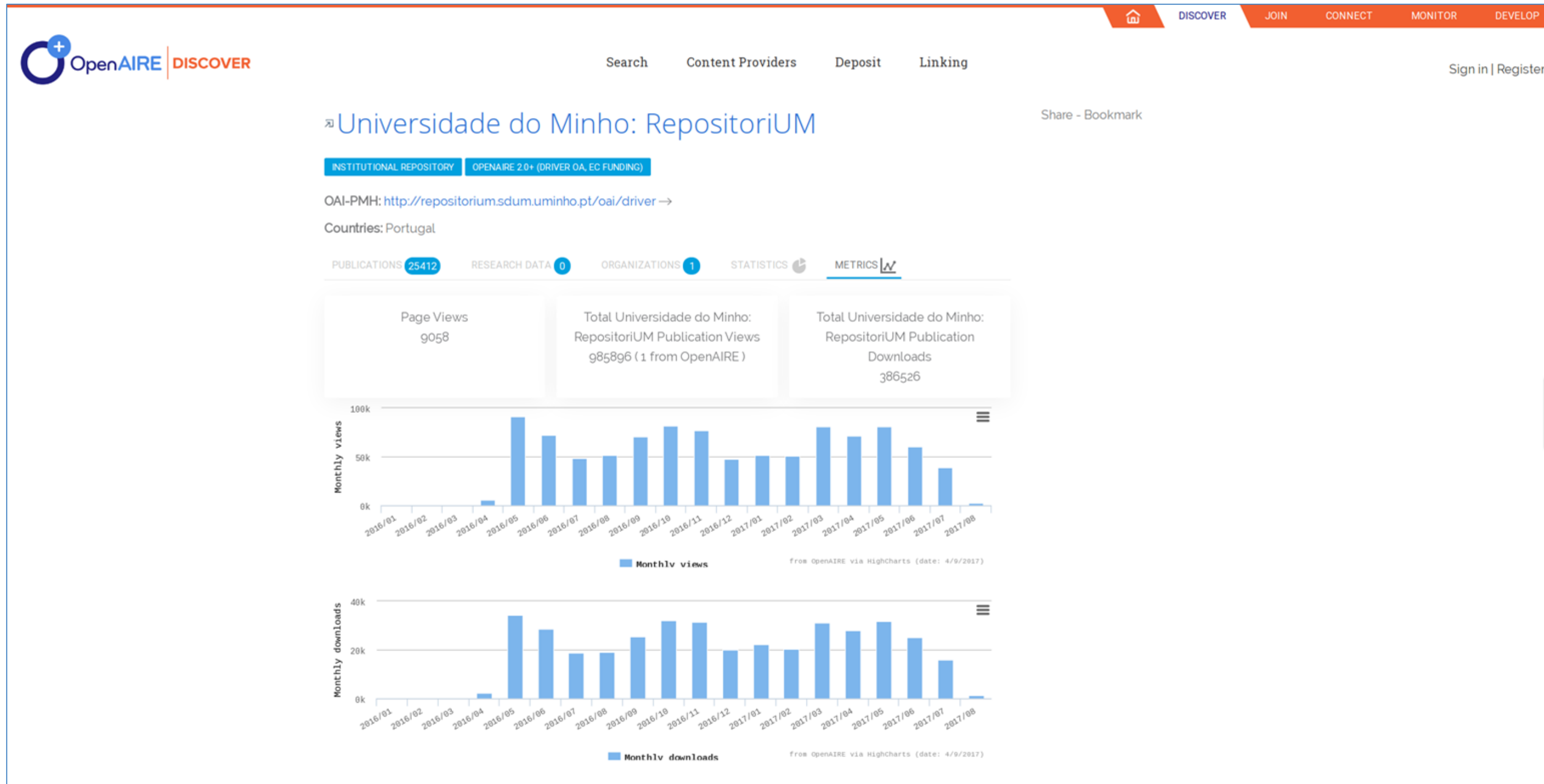
**YOU DON'T HAVE METRICS ENABLED FOR THIS REPOSITORY YET. WOULD YOU LIKE TO ENABLE THEM?**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum ac nulla eget ligula ultrices finibus. Integer eu convallis sapien. Fusce sit amet quam lacus. In dictum, ante sed mattis vulputate, orci dolor suscipit quam, faucibus blandit enim tellus et tellus. Duis quis sem orci. Integer non tortor sed elit sagittis vestibulum. Aenean pulvinar accum ullamcorper, diam sed consequat f

[ENABLE METRICS](#)

```
<!-- Piwik -->
<script type="text/javascript">
  var _paq = _paq || [];
  _paq.push(['enableLinkTracking']);
  (function() {
    var u="//analytics.openaire.eu/";
    _paq.push(['setTrackerUrl', u+'piwik.php']);
    _paq.push(['setSiteId', 49]);
    <% if(handle != null){%>
      _paq.push(['setCustomVariable', 1, 'oaipmhID',"oai:<%= baseUrl %>:<%= handle %>", 'page']);
      _paq.push(['trackPageView']);
    <%>
    var d=document, g=d.createElement('script'), s=d.getElementsByTagName('script')[0];
    g.type='text/javascript'; g.async=true; g.defer=true; g.src=u+'piwik.js'; s.parentNode.insertBefore(g,s);
  })();
</script>
<noscript>
  <p>
    
  </p>
</noscript>
<!-- End Piwik Code -->
```

# Summarized Usage Statistics on the content provider level



# OpenAIRE: A Usage Statistics Hub for Responsible Metrics

- **Quantitative indicators for research**
  - Governance
  - Management
  - Assessment
- **Dimensions**
  - Robust metrics in terms of accuracy and scope;
  - Humble metrics recognizing that quantitative evaluation should support qualitative, expert assessment;
  - Open and Transparent metrics;
  - Diverse metrics by field in order to support the plurality of research and researcher career paths across the system;
  - Reflexible metrics for recognising, anticipating and updating the systemic and potential effects of indicators.



# Open Science as-a-Service

OpenAIRE-connect

*Research Community Dashboard*

*Catch-All-Notification Broker*

# Open Science as-a-Service in the OpenAIRE infrastructure



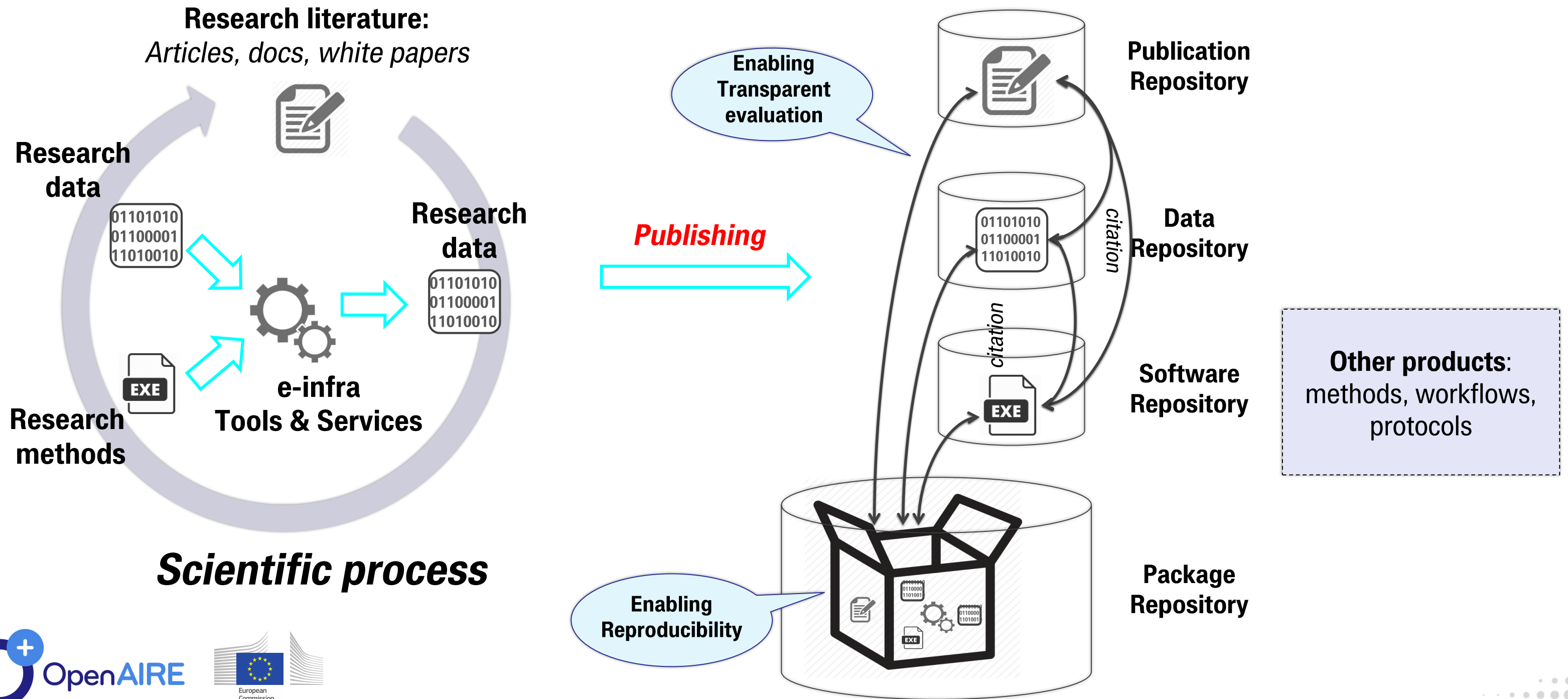
Facilitate **Research Communities** adoption of **Open Science** publishing principles by supporting publishing tools as-a-Service



Facilitate **repositories** at moving towards **Open Science** publishing by supporting notification-based research communication as-a-Service

# Open Science publishing

*Supporting reproducibility and transparent evaluation*



# Open Science Publishing: barriers

## Repositories lack support to Open Science publishing

No support for integration of repositories for software, methods, or packages

Minimal or no support for links between artefacts in different repositories

No support for keeping repositories with up-to-date links between artefacts

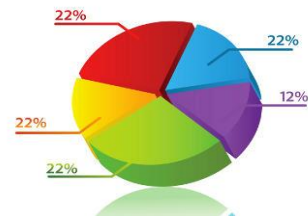
## Research communities lack culture of Open Science publishing

Lack of e-infrastructure and tools for Open Science: e.g. repository limits above, exchange formats, workflows

Difficulties to **self-organize** and **sustain** research communication solutions: e.g. identify the problems, see the benefits, devise solutions, applying economy of scale

# Open Science as-a-Service (OSaaS)

Researchers



Search-Browse-Monitor-  
Research Impact

**Research Community  
Dashboard**

Articles  
Projects  
Data

Software

Methods

Packages

# OSaaS

## OpenAIRE



**Catch-All-Notification  
Broker**

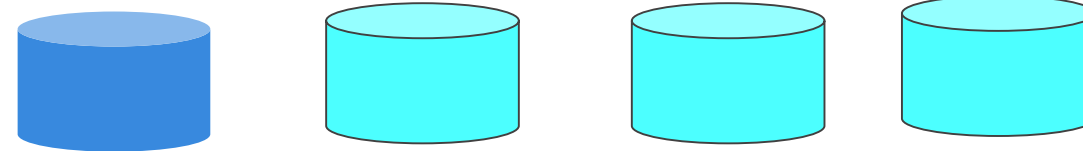
Subscription & Notification

Articles  
Data



Content Providers

Harvesting



Software    Articles    Projects    Data

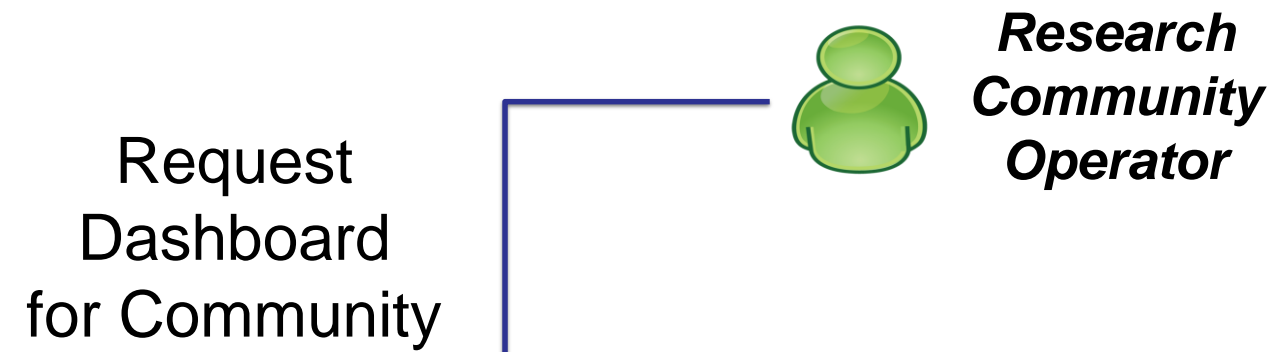


# Research Community Dashboard

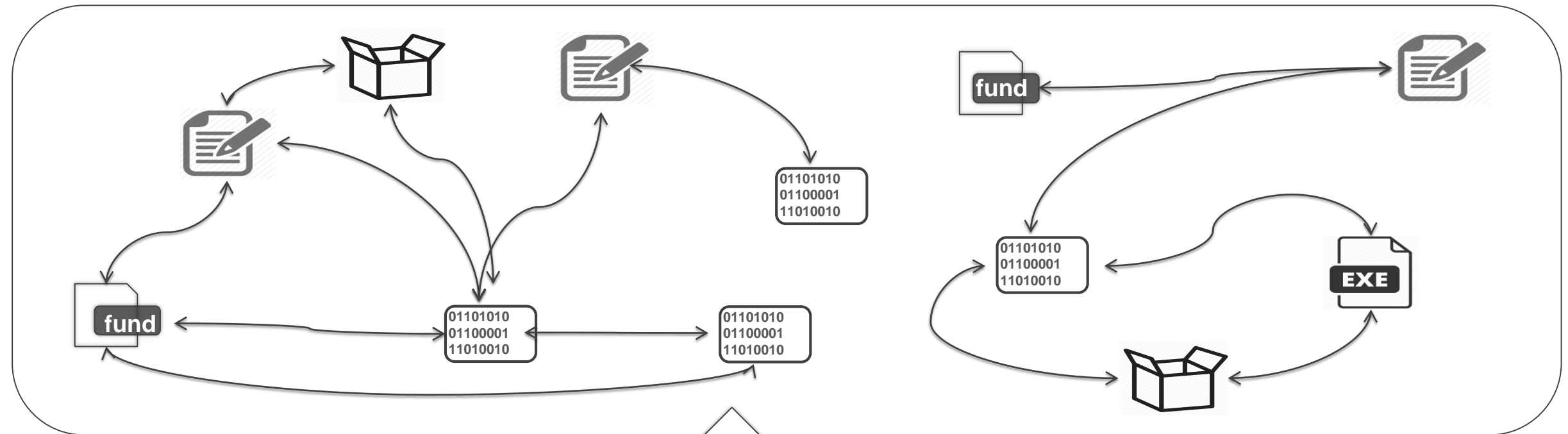
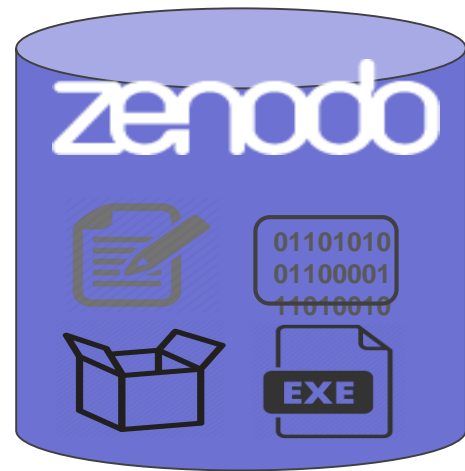
Open Science as-a-Service



# Dashboard for Research Communities



## Research Community Service



De-duplicate

Harvest

Inference

Harmonize



Repositories of publications, datasets, projects, methods, packages

# Dashboard for Research Communities

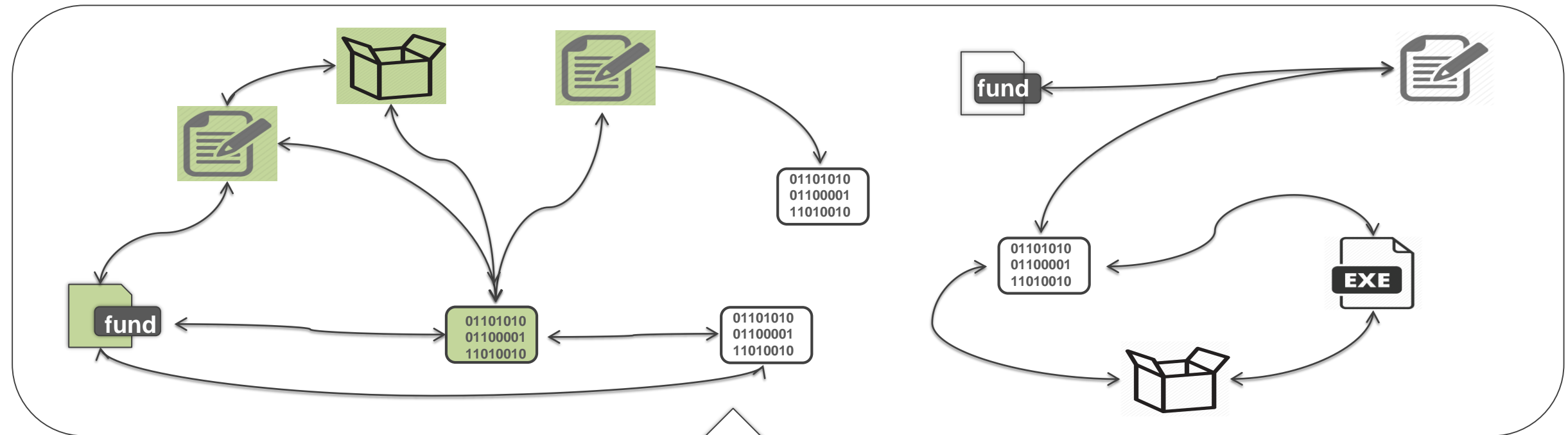
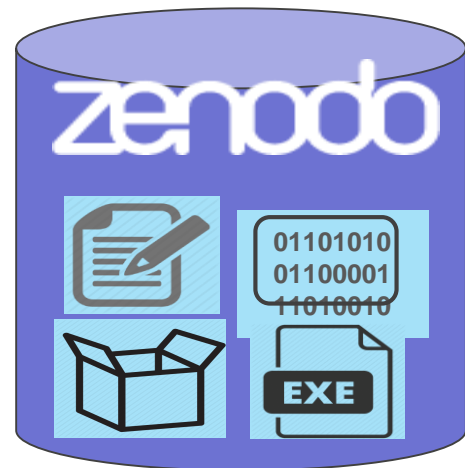


- Deposit (DOI)
- Claim
- Stats: research impact & OA



- Manage users
- Configure stats
- Configure inference

## Research Community Service



**Deduplicate**

**Harvest**

**Inference**

**Harmonize**



**Repositories of publications, datasets, projects, methods, packages**

# Research Communities and Open Science benefits

## Common repository for publishing (deposition) of datasets, methods, and packages

- Can continue their publishing practices, but, if needed they have support for deposition of any artefact

## Collaborative curation of a community-specific research communication domain

- Community information space to share, discovery, and reuse (reproduce) scientific results

## Research impact and statistics

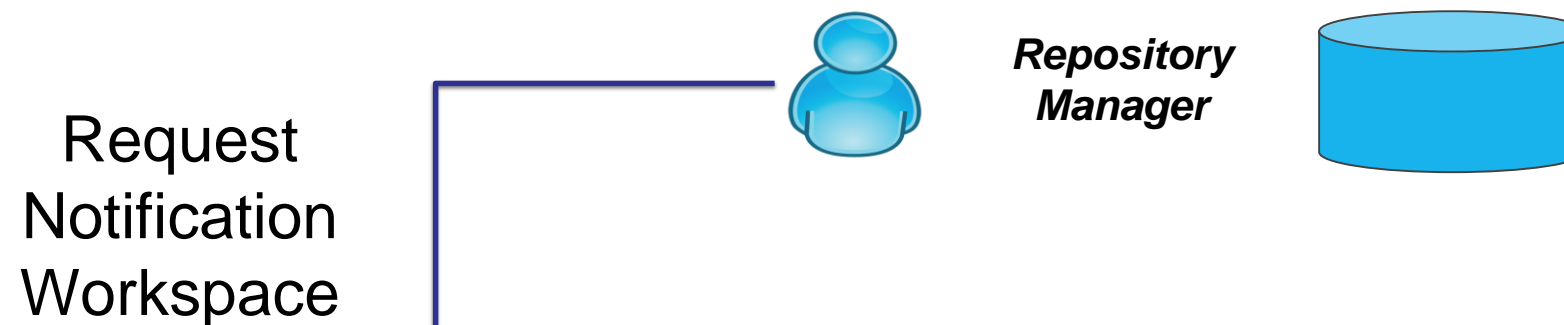
- Scientific reward strategies can be developed

# Catch-All Broker Service

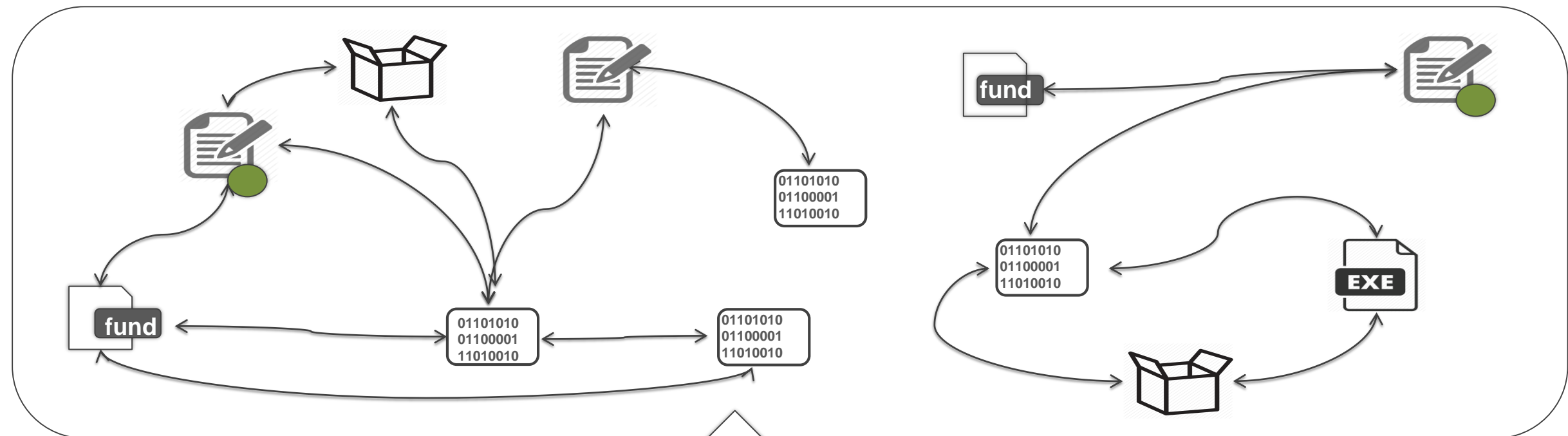
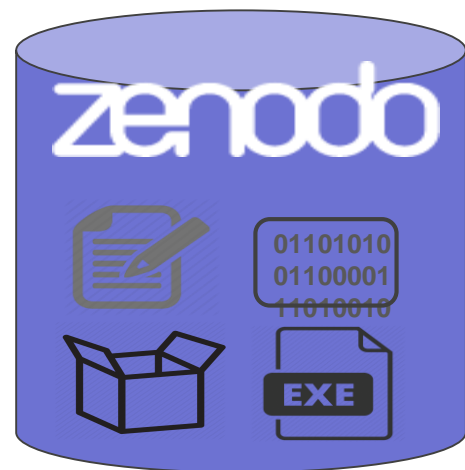
Open Science as-a-Service



# Subscription and Notification for Repositories



## Catch-All Notification Broker Service



De-duplicate  
Harvest

Inference  
Harmonize



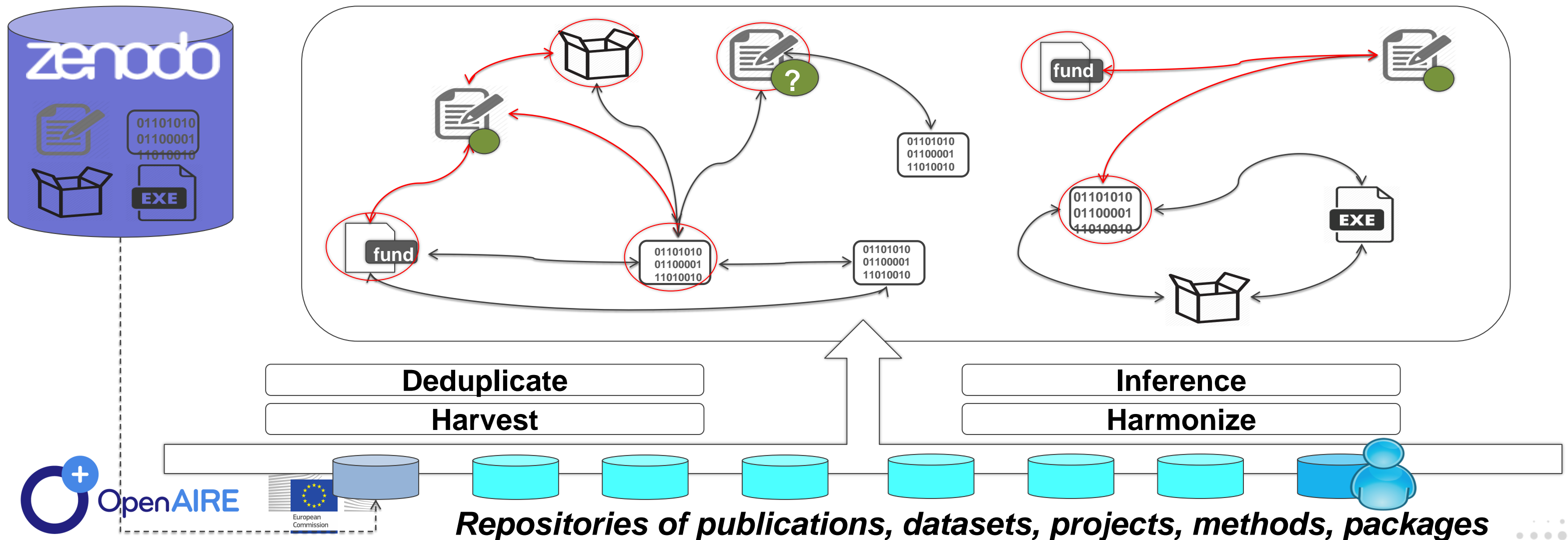
Repositories of publications, datasets, projects, methods, packages

# Subscription and Notification for Repositories

- Subscription: enrichment or additions to repository
- Notification management



## Catch-All Notification Broker Service



# Repositories: Open Science benefits

## Extending repository metadata models to Open Science

- Enabling addition of links to artefacts of any kind

## Keeping their collection up-to-date: enrichments and additions

- “Almost real-time” exchange of information: notifications about links to other artefacts, missing properties, and missing artefacts

## Fostering notification-based and federated dissemination of knowledge

- Enabling repositories to be notified of content of interest, enabling construction of research-focused aggregators by notifications

# OpenAIRE Advancing Open Scholarship

OpenAIRE Advance 2018-2020

*OpenAIRE advance will build on the  
outcomes of the COAR Next  
Generation Repositories*

# Next generation repositories

Repositories forming a distributed global knowledge network have the potential to promote the transformation of the scholarly communication ecosystem.

One of the barriers is a lack of web-based integration with other innovative scholarly services. There is a clear need for repository platforms to adopt modern web-technologies and protocols that will allow them to better interact with more innovative and sophisticated scholarly networked tools and services.

**OpenAIRE advance will build on the outcomes of the COAR Next Generation Repositories Working Group and support the implementation of new functionalities.**

# Next Generation Repositories in OpenAIRE advance

1. *Improving discovery and content transfer.*
2. *Open Metrics for repositories*
3. *Annotation service*

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

 [@openaire\\_eu](https://twitter.com/openaire_eu)

 [facebook.com/groups/openaire](https://facebook.com/groups/openaire)

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