

Eloy Rodrigues

Universidade do Minho
eloy@sdum.uminho.pt

Proyecto y directrices OpenAIRE

Workshop: Interoperabilidad Regional y con OpenAIRE

25, 26, 27 Noviembre – Fundación Oswaldo Cruz, Fiocruz, Río de Janeiro





El Proyecto OpenAIRE



**Las Directrices OpenAIRE: pasado,
presente y futuro**

1



El Proyecto OpenAIRE



**Las Directrices OpenAIRE: pasado,
presente y futuro**



OpenAIRE

OPEN ACCESS INFRASTRUCTURE

for research in Europe

www.openaire.eu

Proyecto piloto en el 7 PM

servicio central en H2020

Science. Set free.

11,456,839 publications and 7,314 datasets from 5,862 repositories and OA journals

OpenAIRE

Dec. 2009 - Nov. 2012

OpenAIREplus

Dec. 2011 – Dec. 2014

OpenAIRE2020

Jan. 2015 a Jun. 2018

OpenAIRE implementa las

POLÍTICAS OPEN ACCESS

en la Unión Europea

OPEN ACCESS EN HORIZONTE 2020

Acceso abierto es el principio general en el nuevo programa marco (CE)

**CADA BENEFICIARIO DEBE
GARANTIZAR EL ACCESO ABIERTO
A TODAS LAS PUBLICACIONES
CIENTÍFICAS CON REVISIÓN POR
PARES RELATIVOS A LOS
RESULTADOS DE LOS PROYECTOS**



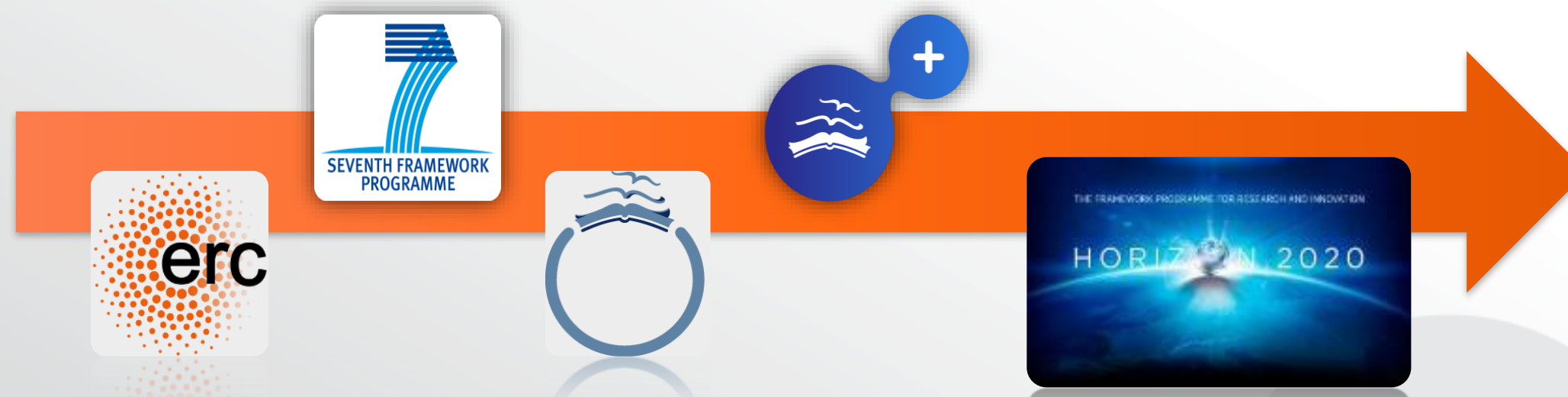
Evolución del mandato OA de la CE

Piloto OA en 7PM (2008)

- Acuerdos con cláusula especial 39
- 20% de las áreas del programa
- Depósito en repositorios
- Directrices OA del ERC: depósito en repositorios (disciplinarias o institucionales) y el período de embargo

Horizonte 2020 (2014)

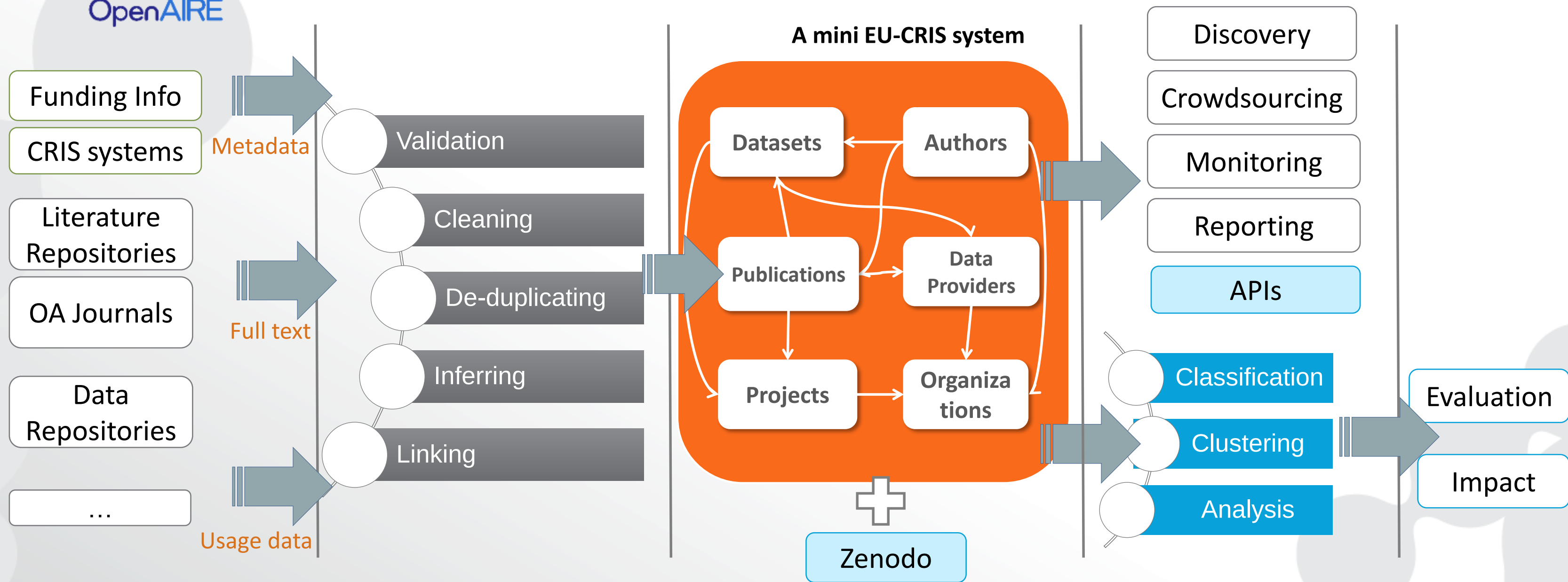
- Todos los acuerdos
- 100% de las áreas del programa
- Depósito en Repositorios
- Proyecto-piloto de Datos Científicos Abiertos



OpenAIRE no es un sistema central

ES EL RESULTADO DE UNA RED DE REPOSITARIOS

**Construido sobre directrices técnicas que
promueven la interoperabilidad
y la visibilidad, impacto... servicios de valor añadido**



Data Providers

OpenAIRE Platform

Services



National &
Regional
Organizations

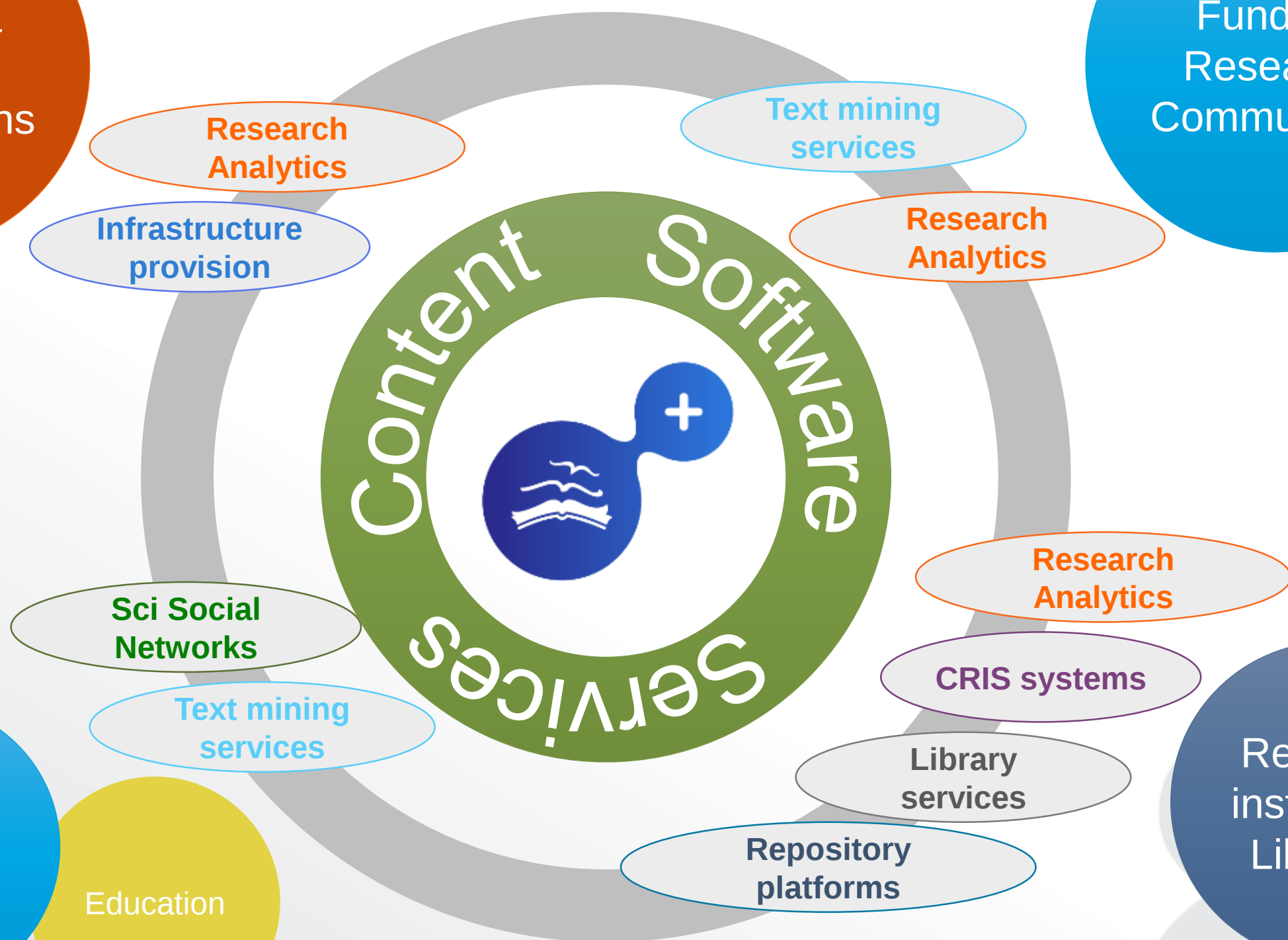
Funders
Research
Communities

Academia

Industry

Education

Research
institutions
Libraries



**SHARE
RESEARCH.
SHAPE A
BETTER
FUTURE.**

Open Access to
research results
benefits everyone.



SCIENCE SET FREE

**Share research outputs.
Shape a better future!**



SUPPORTING THE EC'S
HORIZON 2020 OPEN
ACCESS MANDATE
AND OPEN RESEARCH
DATA PILOT

OpenAIRE is a European **community** and **service-driven infrastructure** that enables Open Science - implementing **Open Access policies** to all research results. It is a **Pan-European Research Information System** with services for discovering, storing, linking, analyzing & monitoring all research outputs across all scientific disciplines.

**50 ORGANISATIONS.
ALL EUROPE'S MEMBER STATES, AND BEYOND.**

Quienes somos

Open Access experts

- Institutional, national and international perspectives on OA policies & e-Infrastructures

Information & Computer Science experts

- Building efficient e-Infra technologies
- State of the art technologies (big data, linked data)

Legal experts

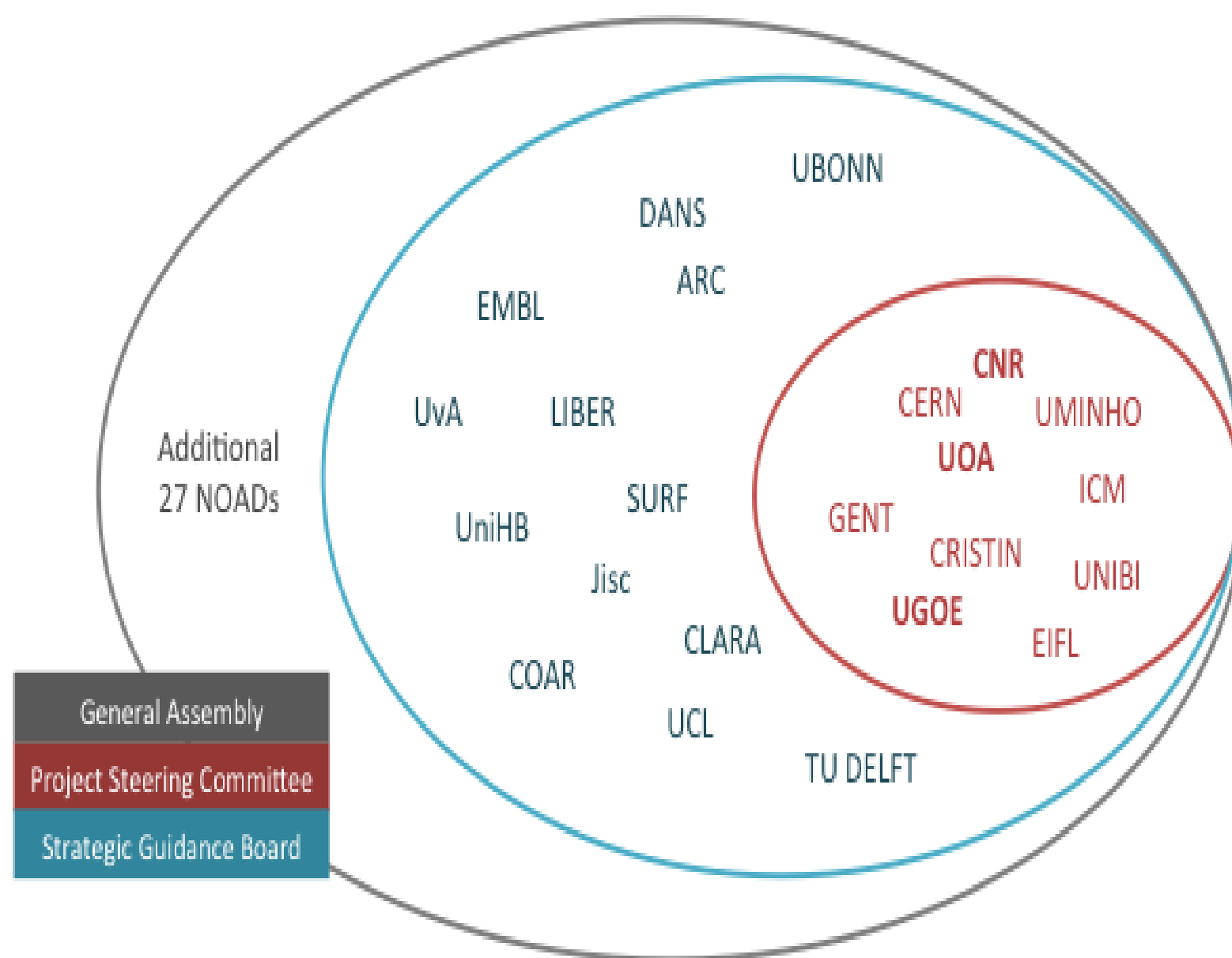
- Legal & policy recommendations

Data communities

- Best practices for data
- Linking to data infrastructures

Consorcio de 50 socios

10 paquetes de trabajo con 40 tareas



ENFOQUE PARTICIPATIVO

**Capitalizando
invertimientos
anteriores**

red humana



e-infraestructura

OpenAIRE2020 Networking Organisation

Operational support
(EIFL)

Strategic coordination (UGOE)

Helpdesk support
(UMinho)

Region North
(CRISTIN)

Region East
(EIFL)

Region South
(UMinho)

Region West
(UGent)

Denmark

Finland

Iceland

Norway

Sweden

Bulgaria

Croatia

Czech Republic

Estonia

Hungary

Latvia

Lithuania

Poland

Romania

Serbia

Slovakia

Slovenia

Cyprus

Greece

Italy

Malta

Portugal

Spain

Turkey

Austria

Belgium

France

Germany

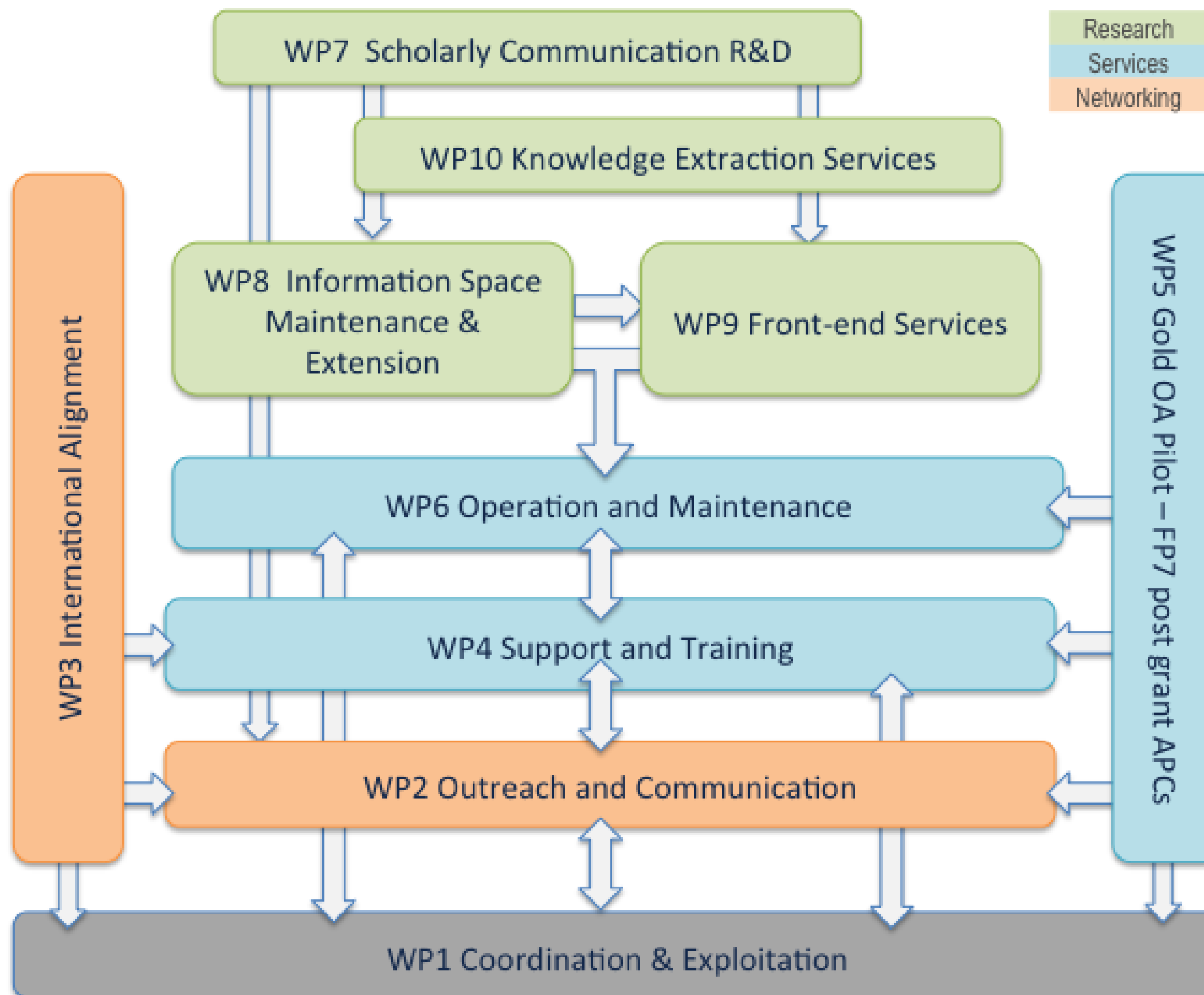
Ireland

Luxemburg

Netherlands

Switzerland

UK



OpenAIRE Portal

Services and information



PARTICIPATE

SEARCH

MONITOR

SUPPORT

OPEN ACCESS



Search in 13,336,838 publications 11,018 datasets from 6,053 repositories and OA journals

The European Open Science Cloud for Research

Leading European initiatives, [EUDAT](#) [LIBER](#), [OpenAIRE](#), [EGI](#) and [GEANT](#) share their joint vision for the European Open Science Cloud for Research which includes eight elements of success for a concrete contribution to the Digital Single Market.

The Open Science Cloud, part of the European Commission's Digital Single Market Strategy, will

RESEARCHERS

Why Open Access. How to comply. What services to use.

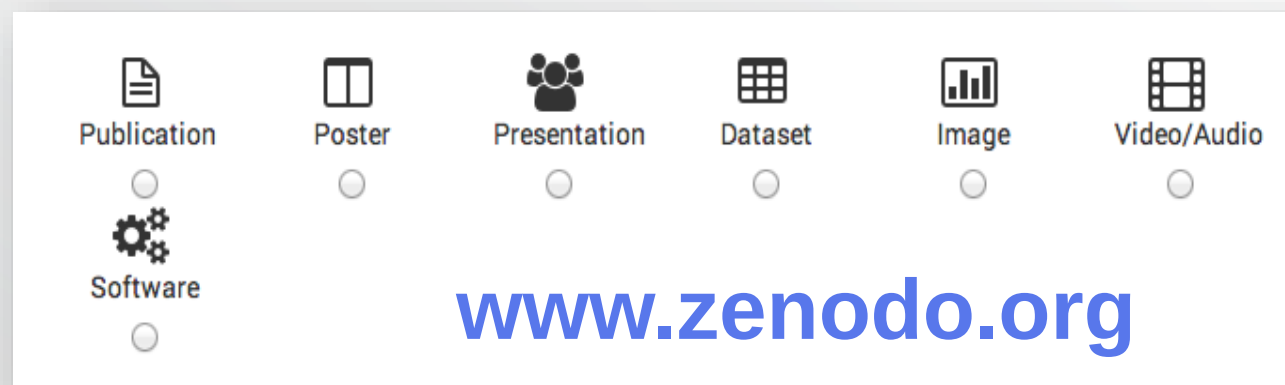
DATA PROVIDERS

How to make your content more visible. What to do to increase quality. How to join.

Zenodo

“Catch-all” repository: OpenAIRE-CERN joint effort

- **Multiple data types**
 - Publications
 - Long tail of research data
- **Citable data (DOI)**
- **Links to funding, pubs, data, software**



API
INTEGRATE YOUR APP VIA
PROGRAMMABLE API.



COMMUNITIES
YOUR DIGITAL REPOSITORY
ON ZENODO.



FUNDING
INTEGRATE INTO
REPORTING FOR RESEARCH
FUNDED BY EUROPEAN COM-
MISSION.



FLEXIBLE LICENSING
NOT EVERYTHING IS UNDER
CREATIVE COMMONS.

H2020: Option to gather, preserve and share
project's scientific output



Aspectos + destacados de OpenAIRE2020

- Mejorar la **interoperabilidad** de todos los recursos del ciclo de investigación
- Apoyo - **mandatos OA H2020**
 - 100% OA en publicaciones científicas
 - Research Data Pilot
- Implementar **Proyecto piloto** FP7 post grant OA publishing funds
- Probar **nuevas formas** de comunicación científica/académica
- Solidificar las **colaboraciones internacionales**
- Establecer **entidad jurídica** OpenAIRE

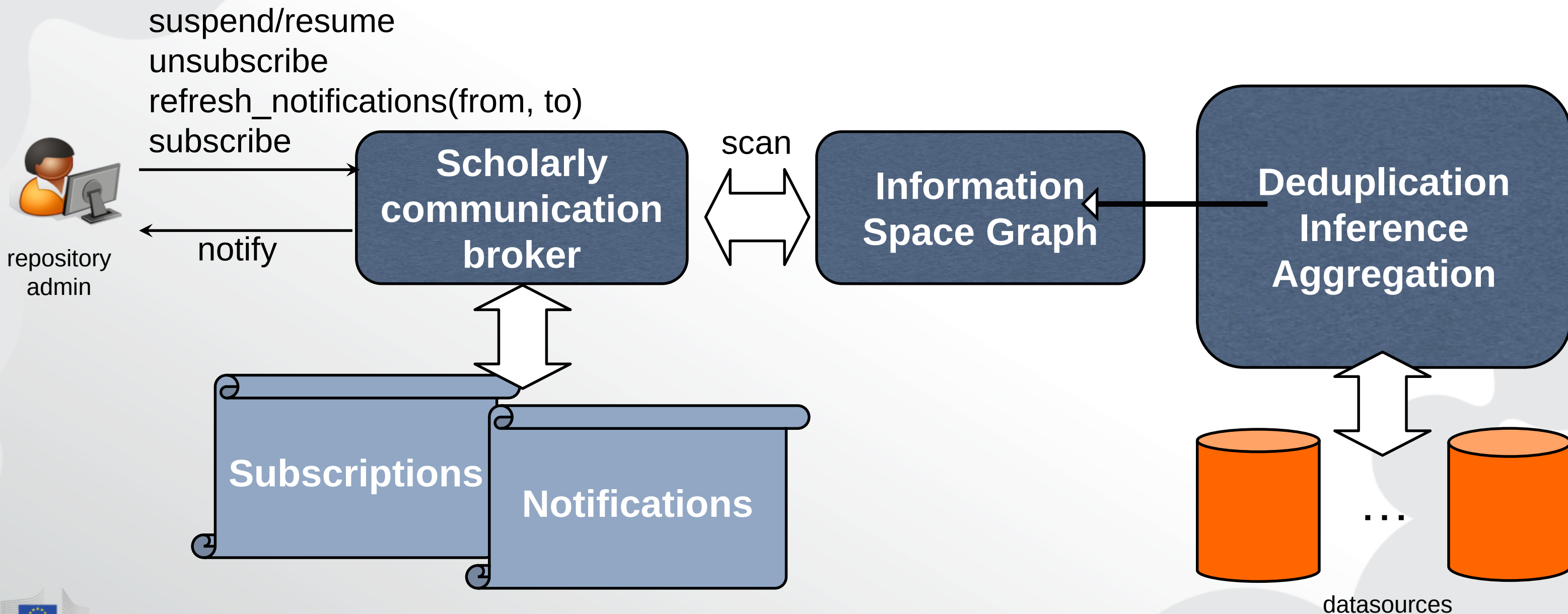
Open Access publication broker

Task 9.4 – CNR, ARC, Jisc, UNIBI, UMINHO, COUPERIN

GOAL

- Institutional repositories may be interested to acquire metadata records of publications that are “**related with**” the repository, i.e. they should/could be part of their collection
 - **Addition:** add to their collection records they were unaware of
 - **Enrichment:** enrich the records they already have with extra metadata information

BROKER: High-level architecture



Examples of additions

- The publication has an author whose organization has a given institutional repository of reference;
- The publication has been funded by a project whose participants (orgs that are beneficiaries of the grant) have a given institutional repository of reference
- The publication has an author with a given institutional repository of reference

Examples of enrichments

- **Open Access status:** e.g. full open access, embargo date
- **Inferred information:** e.g. project fundings, links to datasets, links to publications (citations, similarity), related research initiatives, download statistics, impact measures
- **Standard bibliographic fields:** e.g. persistent identifiers (e.g. DOIs), type of publication, link to files, subject classification

Usage Statistics

Task 8.5 – UNIBI, ARC, UMINHO, COUPERIN, Jisc

- **Align policies and standards for sharing usage data**
 - Sustainable format - learn from Jisc/IRUS experience
- **Collect and analyze usage data to produce metrics**
 - Big data technologies
 - DSpace and other repository platforms plugins
- **APIs to export aggregated usage data**
- **Study legal frameworks for sharing usage data**

Measuring scientific impact

Task 7.2 – CWTS, UoA

- **Goal: move away from Journal Impact Factor**
- **Analysis of multidimensional indicators of scholarly performance based on OpenAIRE data**
- **Investigate indicators for second level impact that link OA to possible social, cultural and economic impact**
 - Patents, EU stats, social media, web impact, etc.

H2020 DATA PILOT

Research Data Management

Support and training

- **Support institutions, PO's, PC's**
 - Educate NOADs on the need for RDM planning and policies
 - Bring out the role of the library
- **Data Management Plan (DMP)**
 - Description, storage, access, validation, versioning, ...
 - Promote good practices through DMP tools (Jisc/DCC, DANS)
- **Promote data repository validation mechanisms (DANS)**

Legal support – two studies

Task 7.1 – UGOE, UvA, ICM

- **Data protection, privacy law, confidentiality**
 - Which data is covered by data protection rules and consequences.
 - Develop legal principles for sharing of personal and sensitive data.
 - Give recommendations for legal and technical measures.
- **Research data under the PSI directive**
 - Evaluate to what extent participants in the Open Research Data Pilot fall under obligations specified by EU and Member States for public-sector bodies (PSBs) on PSI with regard to access and re-use.
 - Come up with exact consequences of those obligations.

Open review to literature

Task 7.4 – UGOE, COUPERIN, CNR

- **What's out there? Systematic analysis of open review systems**
 - Workflows, quality assurance, community involvement, benefits for authors and reviewers, integration with OA e-Infrastructures, ...
- **Prototype for SSH based on hypotheses.org platform**
 - Model the workflow from academic blogs (original articles, reviews about publications, case studies, etc.) for the selection, review and revision of these articles towards peer reviewed publications.
- **Commission two prototypes via open tender calls**

ENTIDAD LEGAL

OpenAIRE to become a legal entity

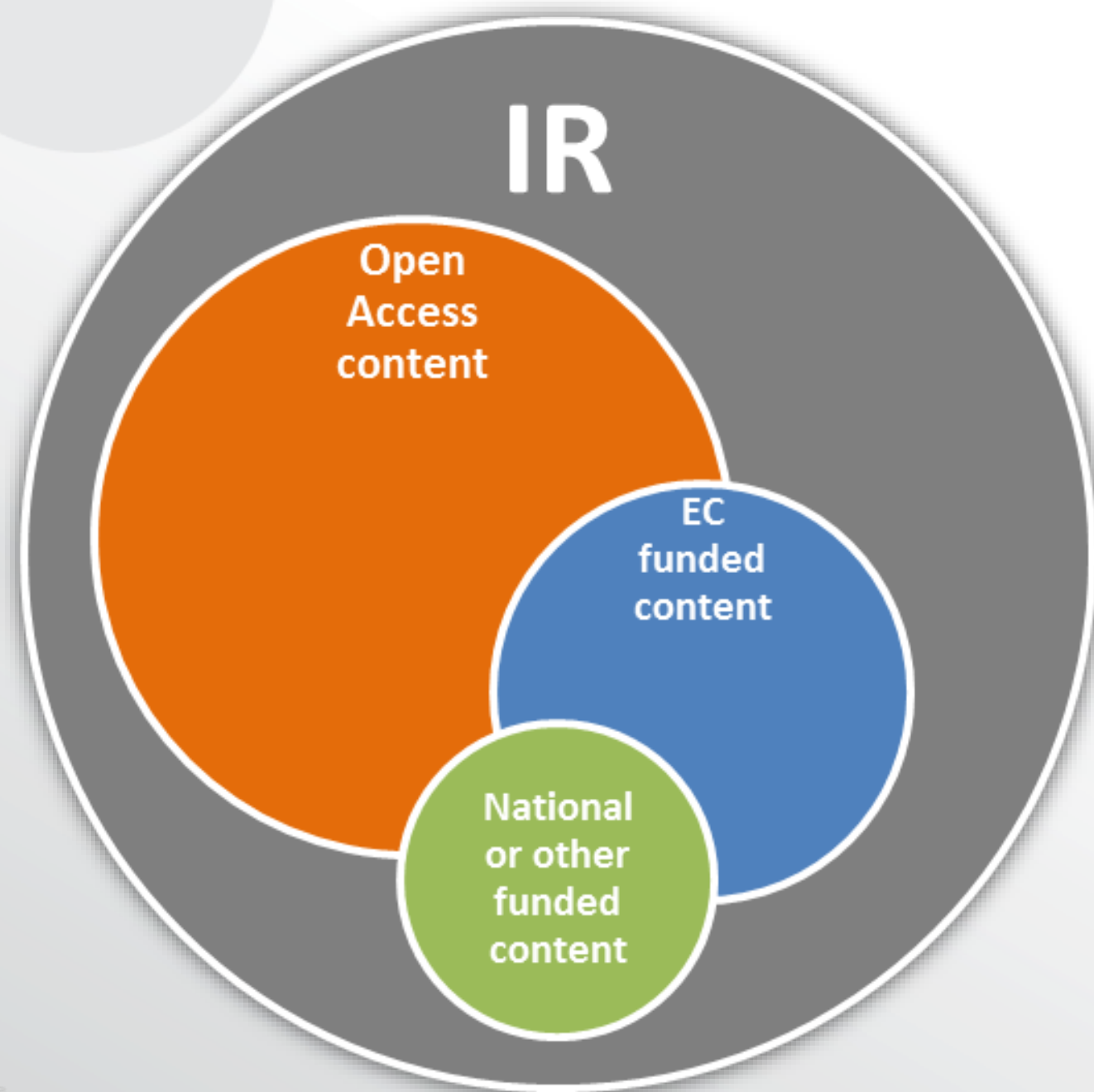


- Early 2016
- Completed
 - Stakeholder analysis
 - Cost benefit analysis
- To be defined
 - Exact mission and objectives
 - Governance model
 - Various (many) logistics

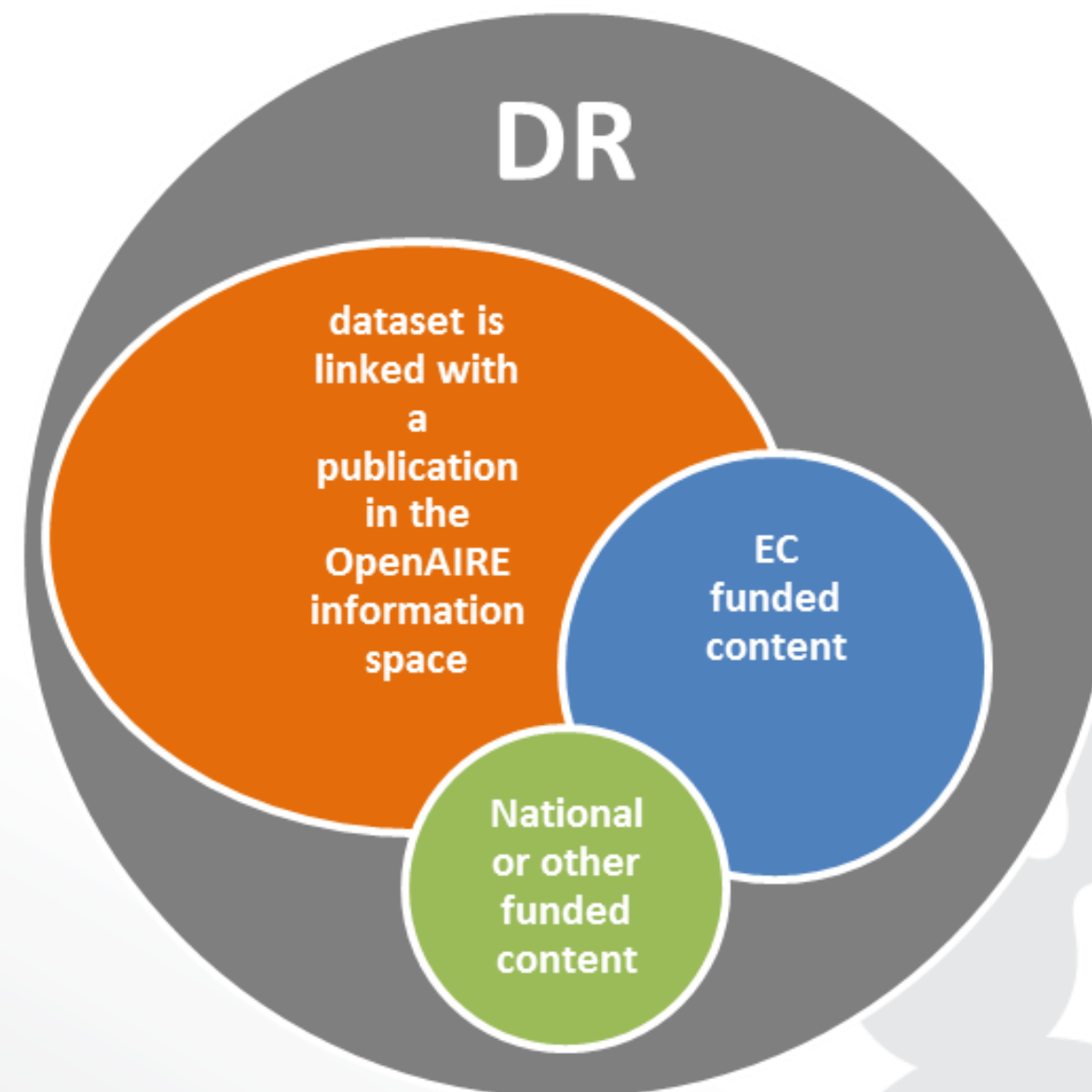
Como funciona OpenAIRE

Política de adquisición de contenidos

Publicaciones en OpenAIRE



Datos en OpenAIRE



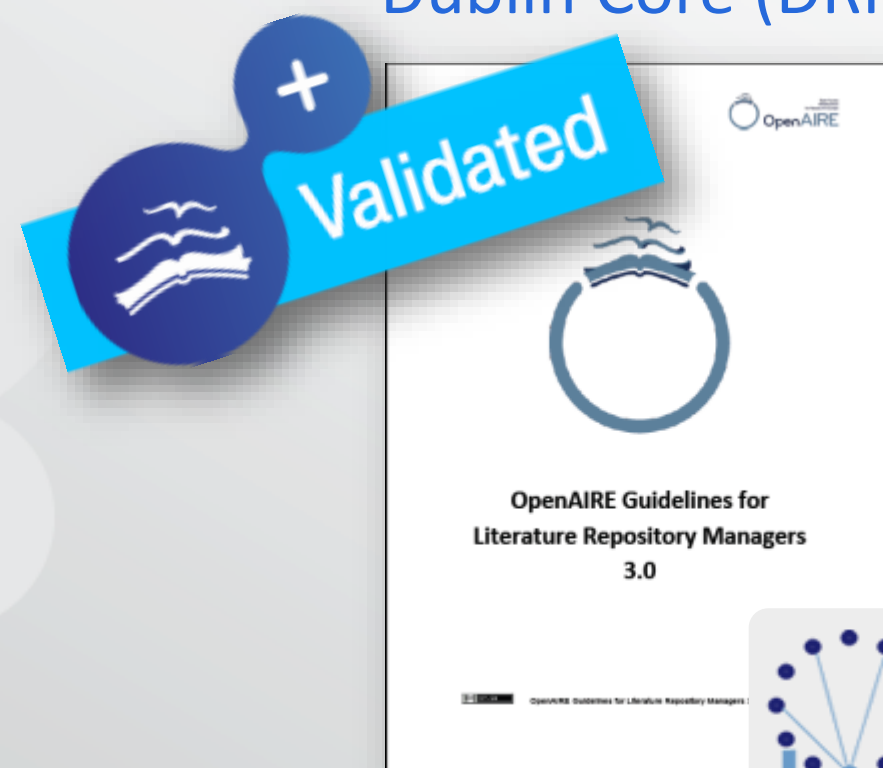
OpenAIRE will expand the current policy to other dataset classes (Open Access datasets) because OpenAIRE would like to have some for quality certification.

INTEROPERABILIDAD

Data providers: 3 directrices OpenAIRE

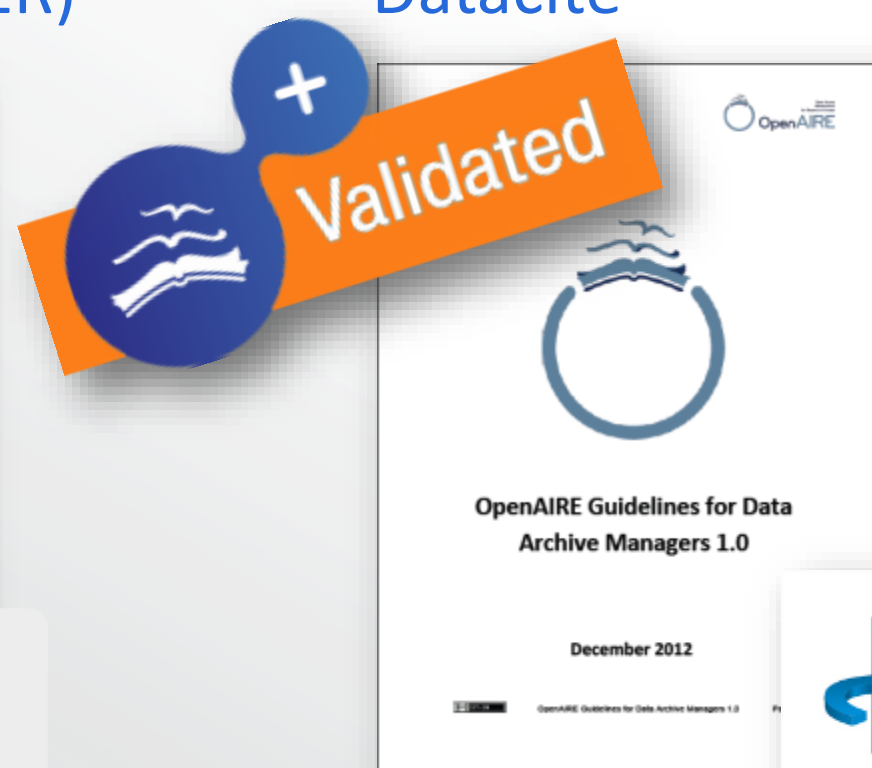
1 Repositorios de literatura

Dublin Core (DRIVER)



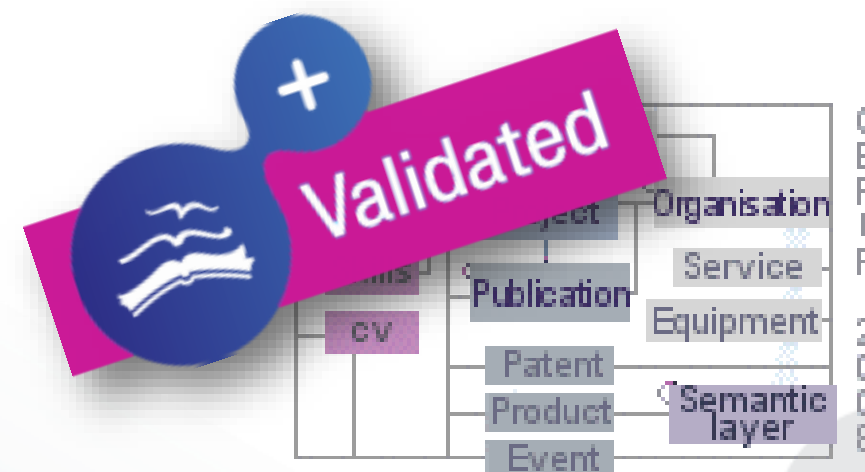
2 Repositorios de datos

Datacite



3 Sistemas CRIS

CERIF-XML

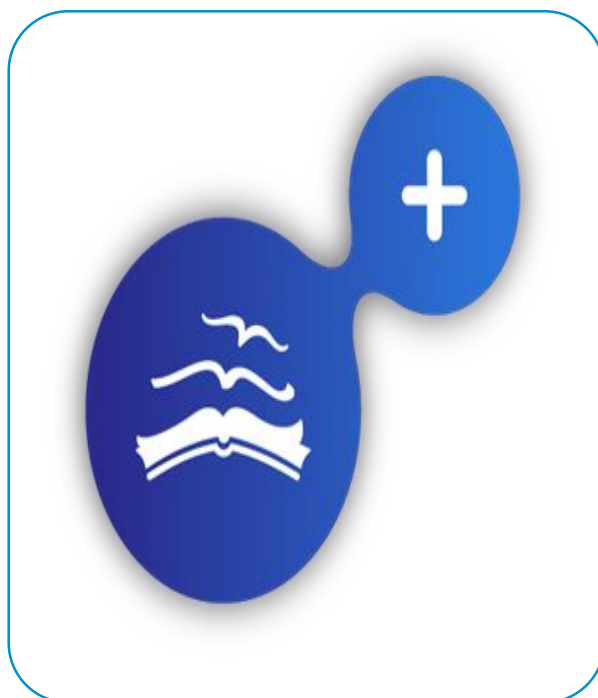


Data Provider and Compatibility (19-Nov-2015)

Data Provider	Quantity	Compatibility
Publication Repositories	522	OA-Basic: 313 OpenAIRE2.0: 19 OpenAIRE2.0+: 115 OpenAIRE3.0: 75
Aggregators/Publishers	16	
Journals	109	
Publication Catalogues	9	
Data Repositories	9	
CRIS	1*	

* Soon (EKT)

2



Proyecto OpenAIRE, hoy y a dos años



Las Directrices OpenAIRE: pasado, presente y futuro

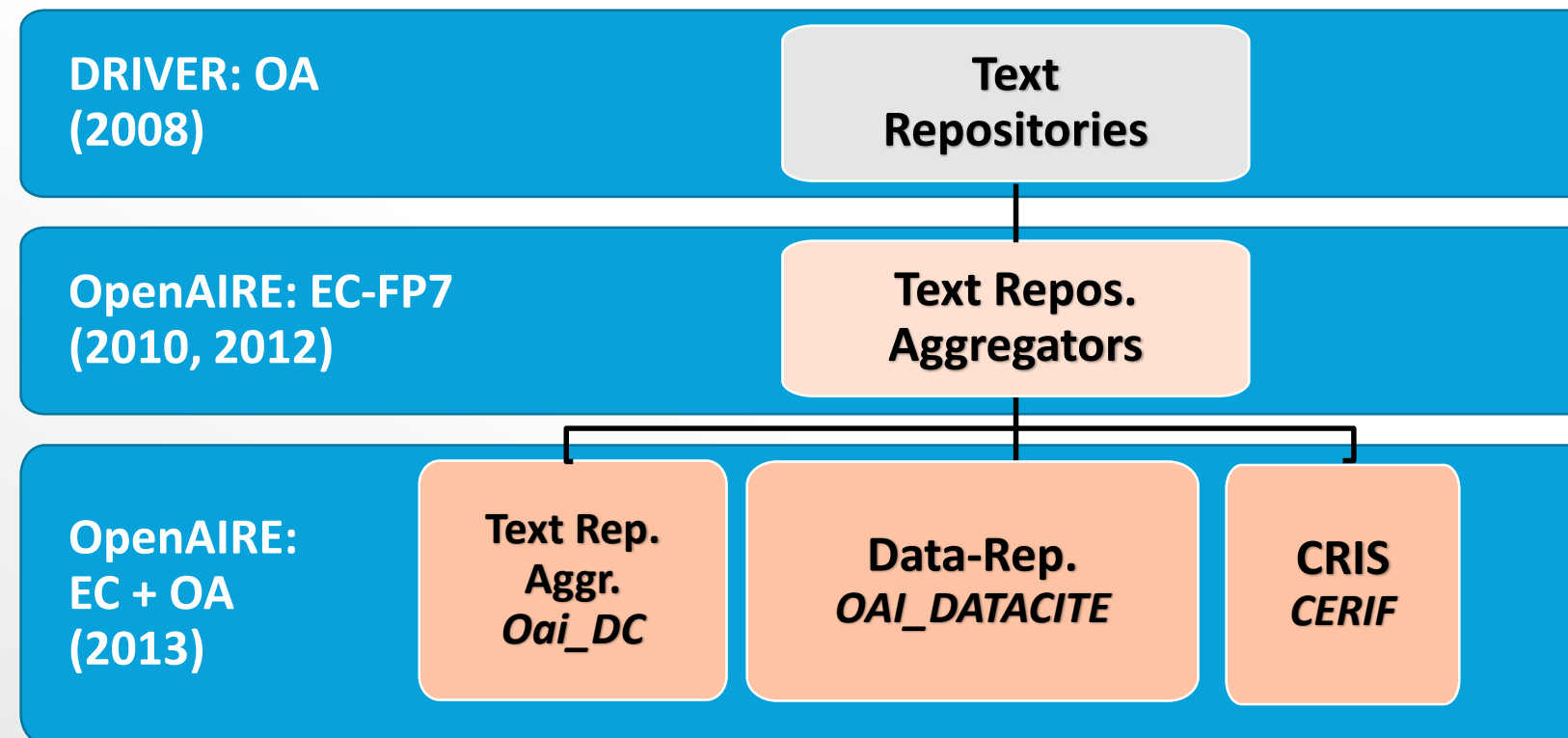
Where we started...

- **'oai_dc' 15 elements based on DCMES schema**
 - As established and 'lowest common denominator' OAI format
- **Used in numerous interpretations -> interoperability issue**
- **Requires agreements on syntax (use of OAI-PMH & oai_dc) and semantics (controlled vocabularies)**
 - DRIVER Guidelines v2 introduced in 2008
 - info:eu-repo Application Profile for encoding schemes
- **Mainly used to describe textual resources**
 - Continued in OpenAIRE Guidelines since 2009

In the meantime...

- **OpenAIRE scope & requirements have expanded since then**
 - **By complementary types of data sources**
 - For research data and research information
 - **Non-textual content**
 - **Publication and data linking**
 - **Project contextualization**
- **New global “authoritative” identifier based infrastructures**
 - **For authors/contributors**
 - **funders / projects**
 - **research outputs**

Guidelines Timeline



3 Guidelines for Data Providers

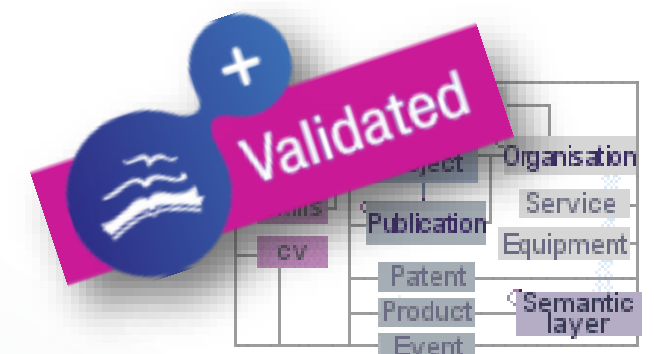
**1 Literature
Repositories**
(and journal platforms)
Dublin Core (DRIVER)



**2 Data
Repositories**
(and archives/data centres)
Datacite



3 CRIS systems
CERIF-XML



euroCRIS
Current Research Information Systems

Directrices OpenAIRE para repositorios

- Continúan las directrices DRIVER v2
- Se establecen en OAI-PMH
- Formato de metadatos: oai_dc
- Application Profile: info:eu-repo
- Con nuevas propiedades para describir:
 - **Funder / Project information**
 - **Embargo Period / Embargo End Date**
 - **Open Access Status / Access Rights**
 - Alternative Identifier
 - Referenced Dataset(s)
 - Referenced Publication(s)

OpenAIRE OAI-Set

- Agrupa metadatos relevantes para OpenAIRE
 - Ver <https://www.openaire.eu/content-acquisition-policy/content-acquisition-policy/content-acquisition-policy>
 - Metadatos sobre Publicaciones en Acceso Abierto
 - Metadatos sobre Publicaciones con financiación EC-FP7 / H2020
 - Metadatos sobre Publicaciones con financiación de otros “funders”

	setName	setSpec*
The OpenAIRE set	OpenAIRE	openaire

projectID

Element name	projectID
DCMI definition	dc:relation
Usage	Mandatory (if applicable)
Usage instruction	A vocabulary of projects is exposed by the OpenAIRE API: http://api.openaire.eu/#cha_projects_http, and available for all repository managers. Values include funder, project name and projectID. The projectID equals the Grant Agreement number, and is defined by the namespace: info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectNumber/Jurisdiction/ProjectName/ProjectAcronym/
Example:	<pre><dc:relation> info:eu-repo/grantAgreement/EC/FP7/123456 </dc:relation></pre> <pre><dc:relation> info:eu-repo/grantAgreement/EC/FP7/12345/EU//Acronym </dc:relation></pre>

ProjectID

namespace

info:eu-repo/grantAgreement/

**Elementos
obligatórios**

Funder/ FundingProgram/ProjectNumber/

EC

H2020

643410

**Elementos
adicionales**

Jurisdiction/ProjectName/ProjectAcronym/

EU

Open Access...

OpenAIRE2020

accessRights

Element name	accessRights
DCMI definition	dc:rights
Usage	Mandatory
Usage instruction	Use values from vocabulary Access Rights at http://purl.org/eu-repo/semantics/#info-eu-repo-AccessRights • info:eu-repo/semantics/closedAccess • info:eu-repo/semantics/embargoedAccess • info:eu-repo/semantics/restrictedAccess • info:eu-repo/semantics/openAccess
Examples:	<pre><dc:rights> info:eu-repo/semantics/openAccess </dc:rights></pre>

embargoEndDate

Element name	embargoEndDate
DCMI definition	dc:date
Usage	Mandatory (if applicable)
Usage instruction	Recommended when accessRights = info:eu-repo/semantics/embargoedAccess The date type is controlled by the name space info:eu-repo/date/embargoEnd/, see http://wiki.surffoundation.nl/display/standards/info-eu-repo/#info-eu-repo-DateTypesandvalue . Encoding of this date should be in the form YYYY-MM-DD (conform ISO 8601).
Examples:	<dc:date> info:eu-repo/date/embargoEnd/2011-05-12 <dc:date>

Alternative Identifier

Element name	Alternative Identifier
DCMI definition	dc:relation
Usage	Recommended
Usage instruction	List alternative identifiers for this publication that are not the primary identifier (repository splash page), e.g., the DOI of publisher's version, the PubMed/arXiv ID. The term is defined by info:eu-repo/semantics/altIdentifier info:eu-repo/semantics/altIdentifier/<scheme>/<identifier> where <scheme> must be one of the following: ark,arxiv, doi, hdl, isbn, purl...
Example	<dc:relation>info:eu-repo/semantics/altIdentifier/doi/10.1234/789.1 </dc:relation>

Referenced Dataset

Element name	Referenced Dataset
DCMI definition	dc:relation
Usage	Recommended
Usage instruction	Encodes links to research datasets connected with this publication. The syntax of info:eu-repo/semantics/dataset is: info:eu-repo/semantics/dataset/<scheme>/<identifier> where <scheme> must be one of the following: ark,arxiv, doi, hdl, isbn, purl...
Example	<pre><dc:relation> info:eu-repo/semantics/dataset/doi/10.1234/789.1 </dc:relation></pre>

Referenced Publication

Element name	Referenced Publication
DCMI definition	dc:relation
Usage	Recommended
Usage instruction	Encode links to publications referenced by this publication. The syntax of info:eu-repo/semantics/reference is: info:eu-repo/semantics/reference/<scheme>/<identifier> where <scheme> must be one of the following: ark, arxiv, doi, hdl, isbn...
Examples:	<pre><dc:relation> info:eu-repo/semantics/reference/doi/10.1234/789.1 </dc:relation></pre>

Guidelines for Research Data based Data Sources

- **Based on DataCite metadata scheme**
 - Proven standard for heterogeneous data sources and cross-disciplinary archives
 - Maintained by a trusted and sustainable organization
 - Support of data citation
- **Metadata Format: oai_datacite**
- **OAI-Set: openaire_data**
- **Zenodo as reference implementation for OpenAIRE Data Repository**

Guidelines

Guidelines for Research Data based Data Sources

- Adaptations for OpenAIRE:
 - Incl. additional PID schemes to DOI, e.g.URN, ARC ...
 - Recommends links to related publications & datasets
 - Recommends Contributor prop.
to relate dataset to funding information
 - Date and Description prop. are mandatory
 - Enforces (info:eu-repo) encoding scheme
on Rights property

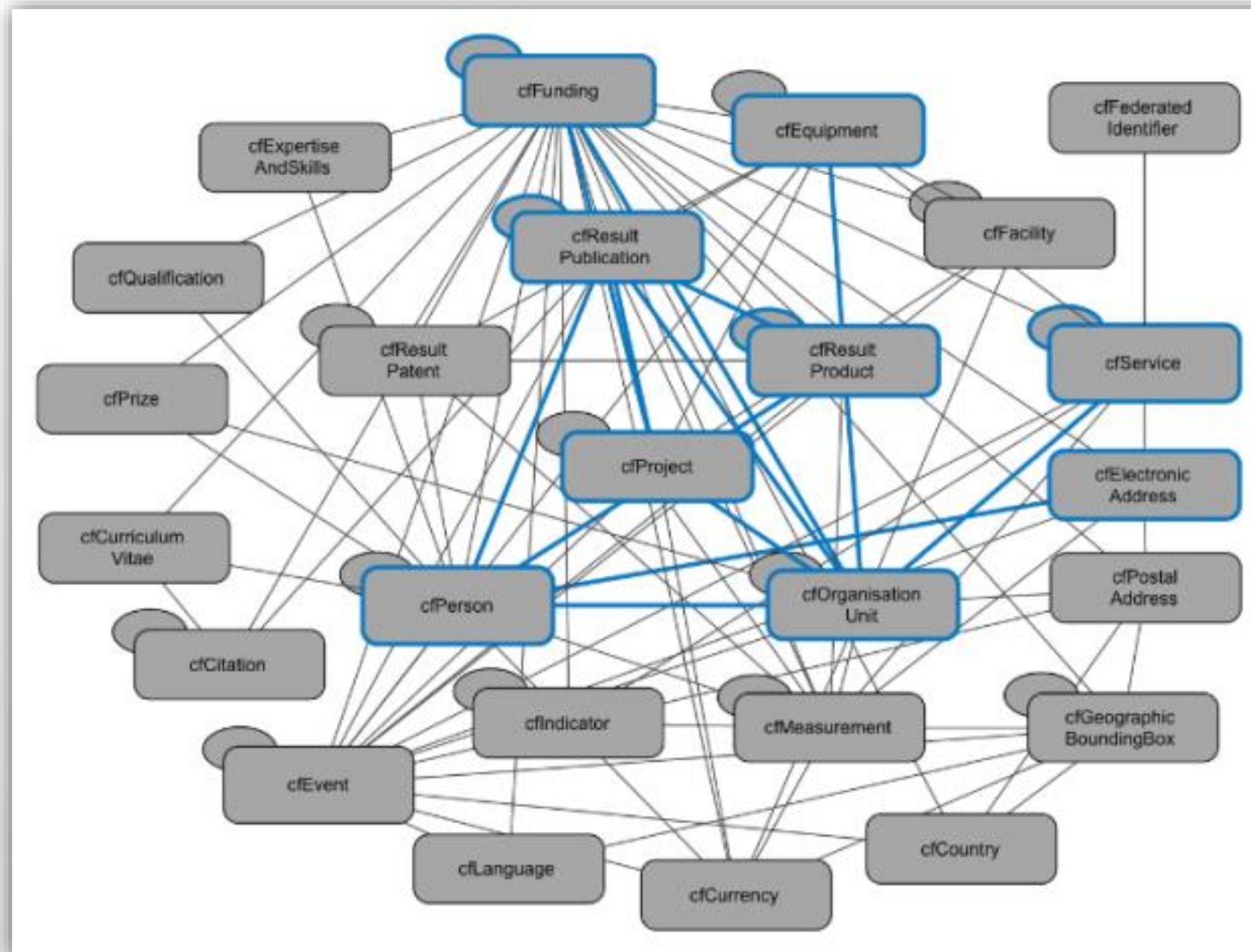
Guidelines for CRIS Managers based on CERIF-XML

- CERIF as de facto standard for CRIS data model
 - Entities for organization, person, project, output, data ..
- CERIF-XML as data exchange standard
(in the CRIS world)
- Maintained by euroCRIS as sustainable organization

Guidelines for CRIS Managers based on CERIF-XML

- **Adaptations for OpenAIRE:**
 - Subset of CERIF entities relevant for OpenAIRE
- **Semantic Interoperability**
 - Reuse of CERIF Semantics vocabularies for classification and definition of relationships between entities
 - Consistency with the other OpenAIRE Guidelines
 - E.g. resource types, access rights

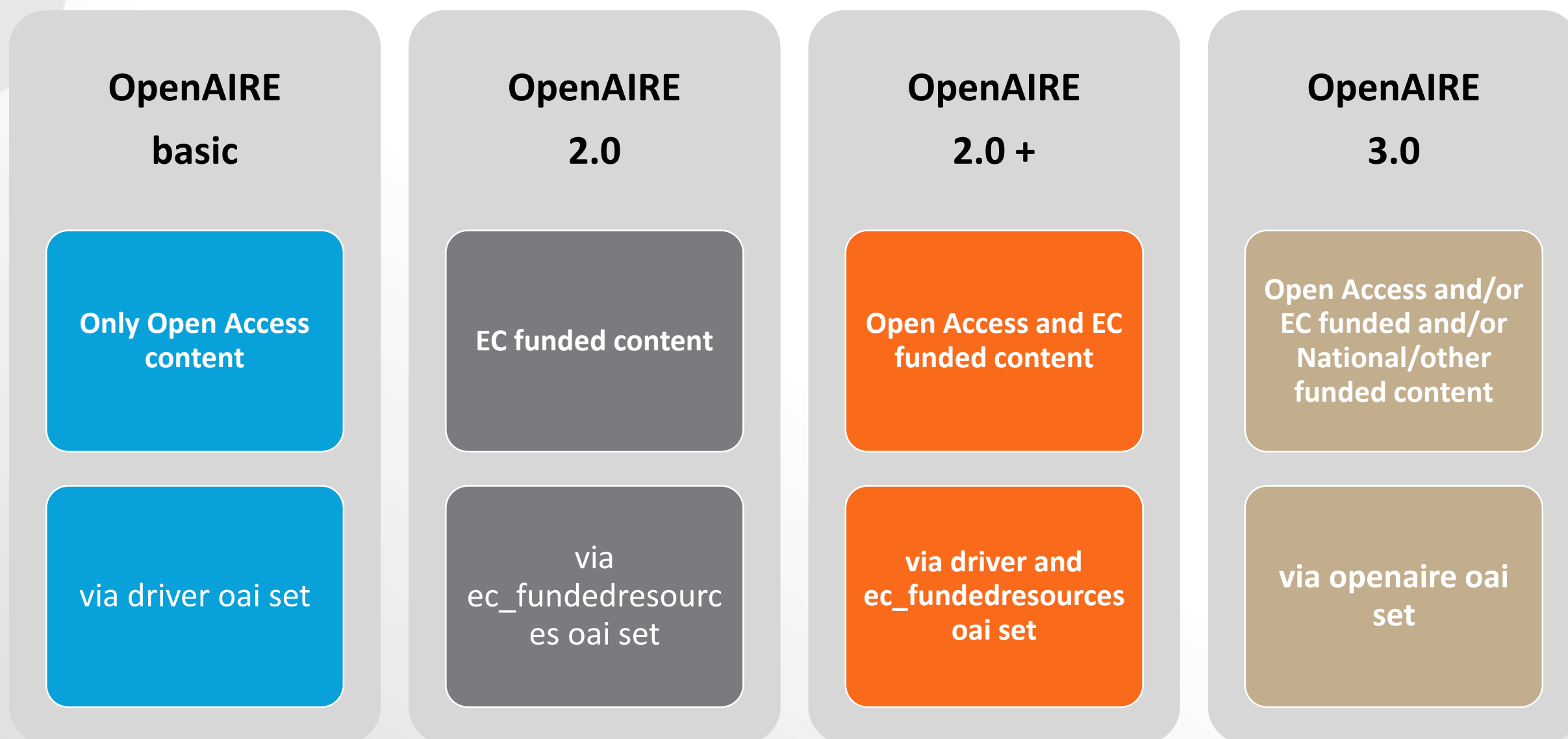
Relevant CERIF Entities for OpenAIRE



Guidelines for CRIS Managers based on CERIF-XML

- **Syntactic Interoperability**
 - Distinct namespace for OpenAIRE CERIF schema
 - Nesting only of multi-lingual attributes, federated identifiers, linked entities
- **System Interoperability:**
 - Transfer Protocol: OAI-PMH
 - OAI-PMH sets for
 - Each type of entity
 - The entire CERIF graph

Compatibilidad OpenAIRE : Niveles y OAI Sets



El futuro de las Directrices OpenAIRE

- **Cuestiones y necesidades identificadas**
 - Utilización de identificadores (de autores, financiadores, organizaciones, etc.)
 - Utilización de vocabularios controlados (tipo de recurso, tipo de acceso, clasificación de asuntos, versión, etc.)
 - Información (citación) bibliográfica detallada (por ejemplo, título revista, ISSN, volumen, número, página inicial y final)
 - Expresar para donde resuelve un URI (para una “landing page” o para un o varios ficheros de contenido?)
 - Utilizar/suportar diferentes formatos de “serialización” y metadatos(para allá de OAI-DC simples):
 - Dcterm based app-profile (as XML, as RDF)?
 - Datacite metadata format?
 - BibJSON?

El futuro de las Directrices OpenAIRE

- **Cuestiones prioritarias**
 - Utilización de identificadores (de autores, financiadores, organizaciones, etc.)
 - Utilización de vocabularios controlados (tipo de recurso, tipo de acceso, clasificación de asuntos, versión, etc.)

Opinión y participación de La Referencia

- **Opinión y sugerencias sobre**
 - Futuro Namespace
 - Solución para los authorIds e ProjectIds
 - Derechos
 - Vocabulario Versión
 - Otras prioridades?

Gracias! Al trabajo!

 www.openaire.eu

 @openaire_eu

 facebook.com/groups/openaire

 linkedin.com/groups/OpenAIRE-3893548

