



OpenAIRE

Open Access Infrastructure for Research in Europe

The OpenAIRE Infrastructure

EC Policy on Open Access and the OpenAIRE Initiative
EGI Scientific Publications Repository *Workshop*

Pasquale Pagano

CNR - ISTI

*Courtesy by Donatella Castelli,
Yannis Ionnadis, Paolo Manghi,
and Natalia Manola*



OpenAIREplus objectives

- **Promoting Open Access:** best practices and business models for articles and datasets
- **Providing services:** aggregation and linking of articles, datasets, projects, funding schemes (EC and national)
- **Measuring impact:** any relevant entity, SC39 policy



EC Open Access policies

- **2009:** FP7 OA Pilot and ERC OA guidelines
 - 20% of FP7 in seven scientific areas
- **2012:** EC Communication “Towards better access to scientific information: Boosting the benefits of public investments in research” (17th July)
 - Open Access of all publications in H2020
 - Pilot scheme on open access to and re-use of research data generated by projects in selected areas of H2020.
- **2012:** EC Recommendation on access to and preservation of scientific information (17th July)
 - Member states should define clear policies for Open Access to results of publicly funded research (publication & data)



Open Access

It is not *if* Open Access

It is *how* Open Access

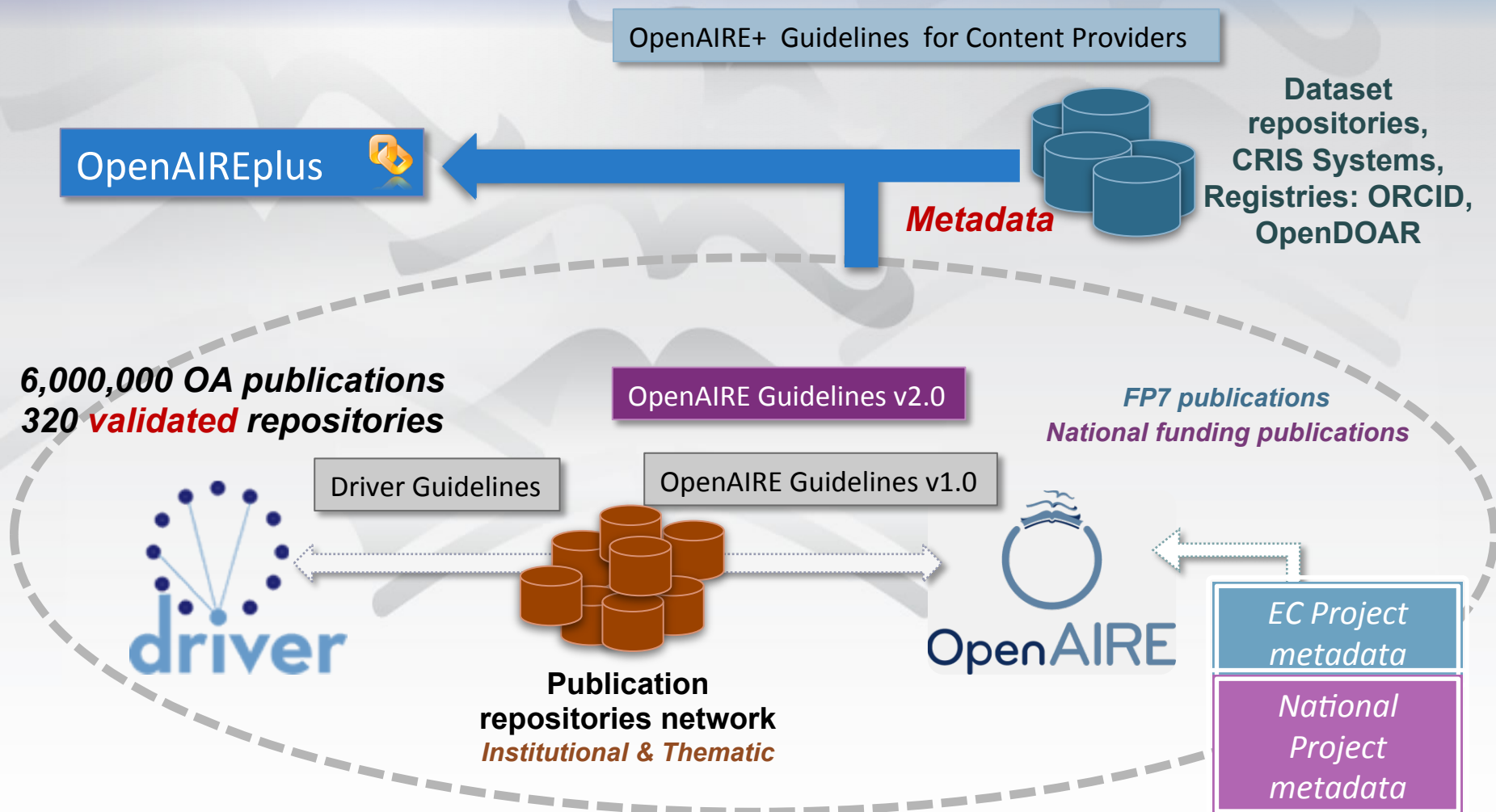


What is OpenAIRE?

The European Open Access
data infrastructure
for
Scholarly & Scientific Communication



Past, Present, OpenAIRE





Supporting OA policy in Europe

- Technical infrastructure *not enough*
- Network of National Open Access Desks
 - Think global, act local – *all member states*
 - Diverse national research environments
 - Reach out to researchers, organizations, infras



OpenAIRE infrastructure

Human Infrastructure

Networking for Open Access
Community building,
dissemination, liaisons, studies

Data Infrastructure

Research and Services
Data collection, aggregation,
curation, enrichment, provision



Human Infrastructure



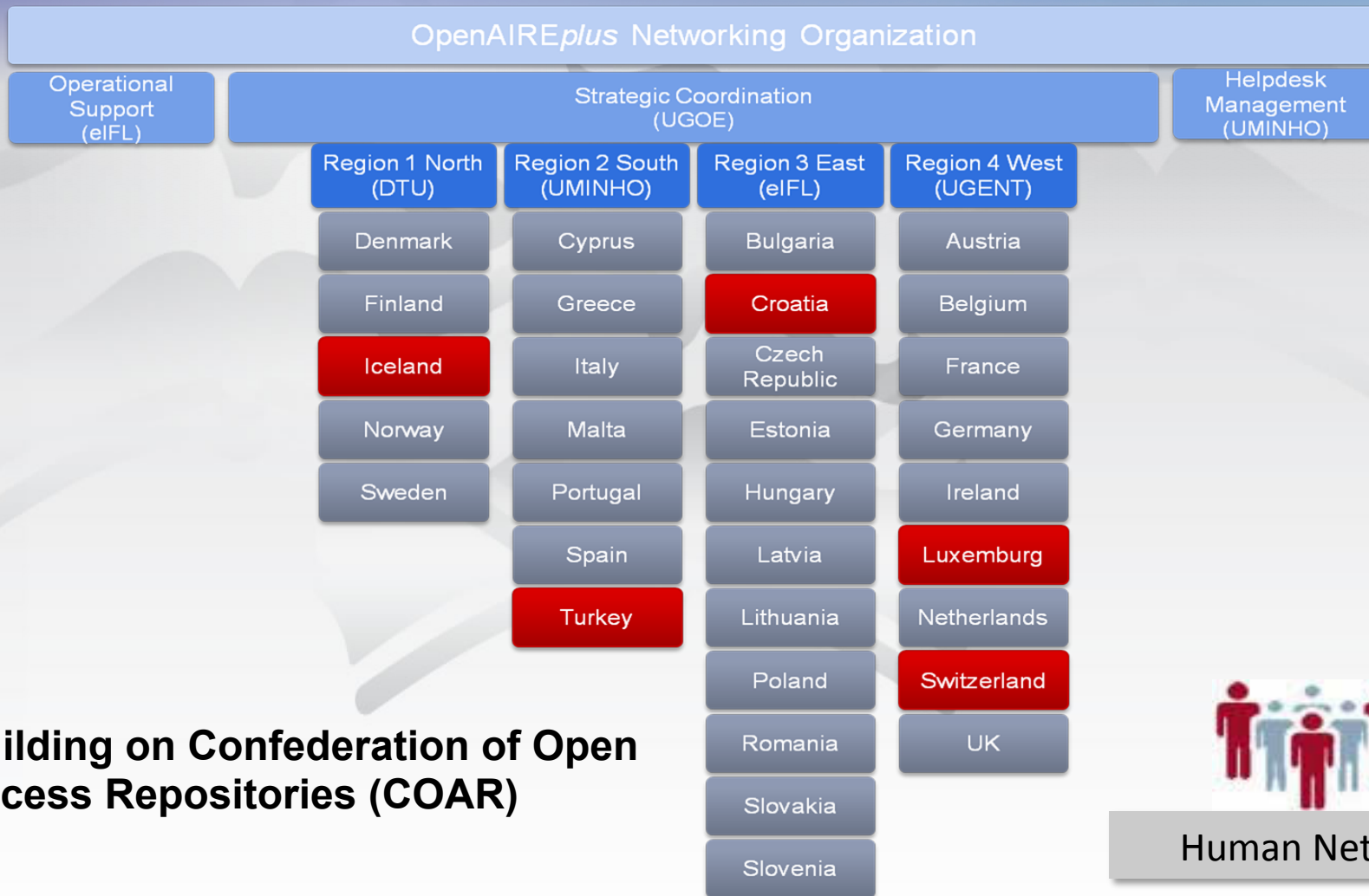
Main Goals and Activities

A European network of people sharing experiences & best practices

- **OA Helpdesk** for European researchers
 - Dissemination, liaison, community building
- National and European **scholarly publication and data infrastructure initiatives** alignment on OA
 - DataCite, ESFRI infra projects (CLARIN, DARIAH), EuroCRIS (CERIF), EUDAT, Europeana, ORCID, ...
- **Guidelines** for content providers
- **IPR issues** for pubs and research data
- **Feasibility/financing models** for OA



National Open Access Desks





Guidelines for Content Providers



Repositories



- Extending OpenAIRE Guidelines for repository managers: OAI-PMH protocol: DC + project encoding + license encoding



Data archives

- DataCite metadata



CRIS systems

- CERIF metadata



Entity registries

- Ad-hoc solutions (e.g., OpenDOAR, CORDA, ORCID)



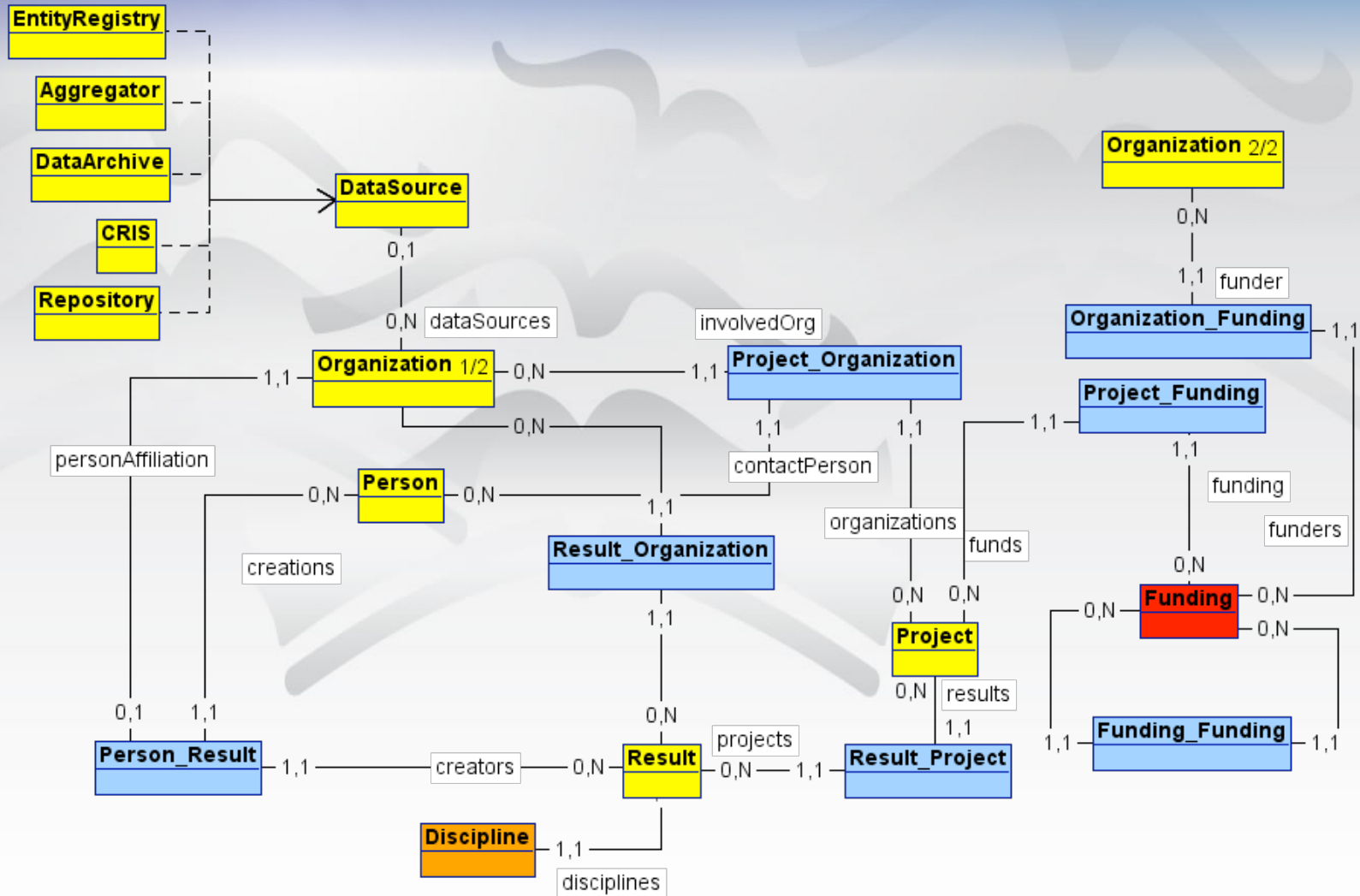
Data Infrastructure

OpenAIReplus Data Infrastructure *Content-wise*

- **Data model:** Linking **publications (OA or not), datasets, projects, persons, organizations, funding** schemes and agencies, ...
- **Data sources:** publication repositories (plus 6,000,000 OA pubs indexed by the **DRIVER data infra**), data archives, entity registries (CORDA, OpenDOAR, ORCID), CRIS systems
- **Orphan repository:** Deposition of publication/datasets for authors w/o a repository of reference

+ OpenAIReplus Data Model

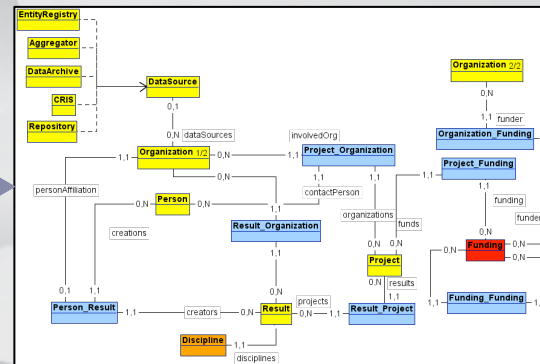
CERIF&DataCite inspiration



applications to (i) link metadata to projects (ii) correct/enrich information space



End-user Feedback



(Semi-)Automated provision:

*facilitated by “mandates” typically
issued by funding agencies
(e.g., project reference embedded in publication text)*



Supported by compliance
to "Guidelines
for Content
Providers"

Repositories, archives, entity registries and CRISs

+ Measuring research impact

“xyz centric” (e.g., xyz=infrastructure)

Manual provision:

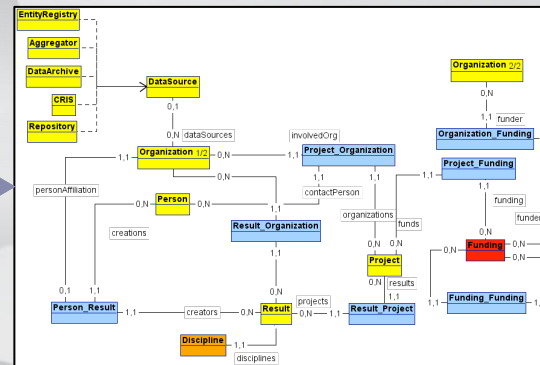
applications to (i) link metadata to projects (ii) correct/enrich information space

publication/dataset & XXX

“Claim”

Any relationships & entities

End-user Feedback



Enriched Data Model (i.e., special tags)

Any relationships & entities

Inference by mining

(Semi-)Automated provision:
facilitated by “mandates” typically issued by funding agencies
(e.g., XXX reference embedded in publication text)

Any relationships & entities

Collection

Automated provision:
Supported by compliance to “Guidelines for Content Providers”

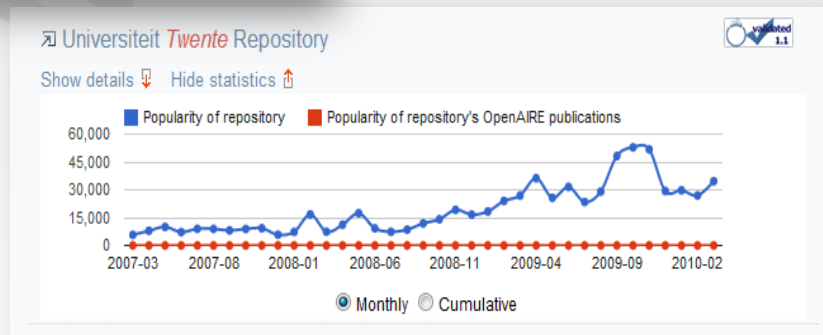
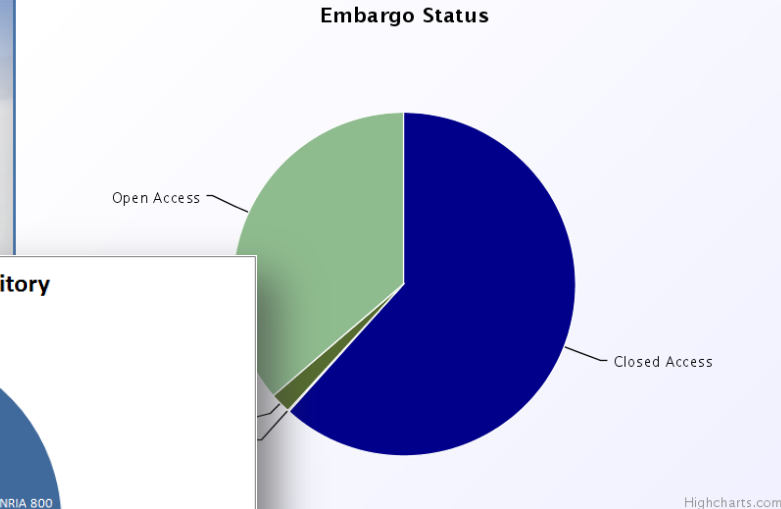
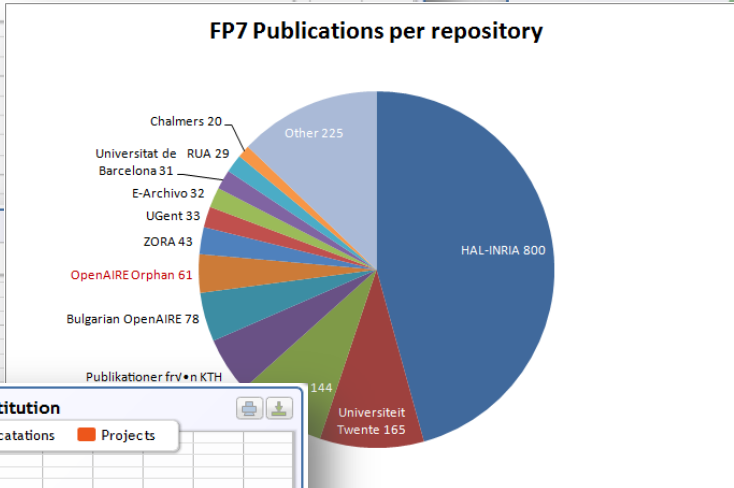
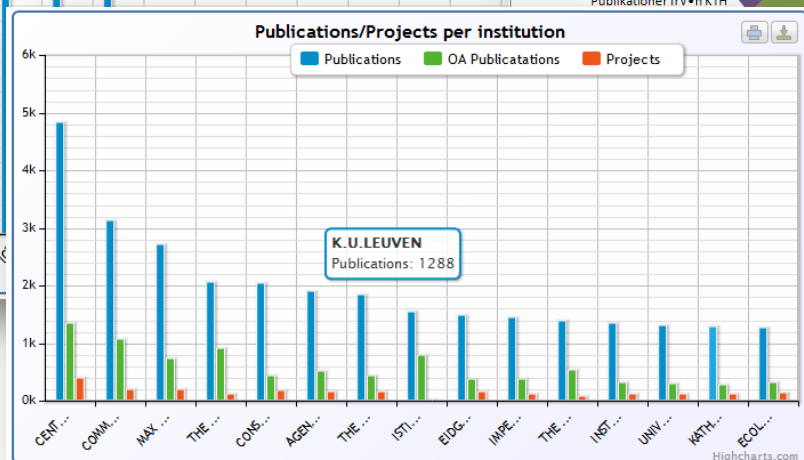
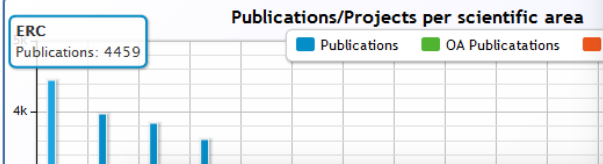
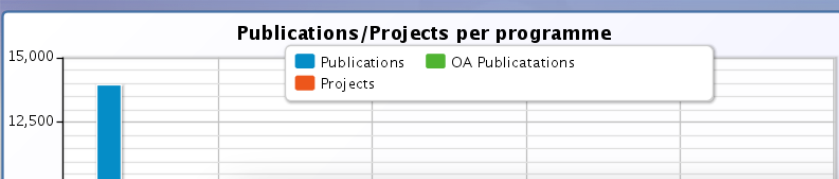
Repositories, archives, entity registries and CRISs

OpenAIReplus Data Infrastructure *Service-wise*

End-users, content managers, data curators, 3rd-party apps

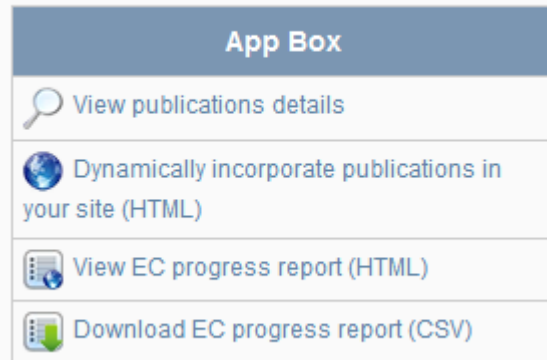
- **(Meta)Data mgmt:** harvesting, aggregation, deposition, claiming, validation, access (portal and APIs)
- **Data curation:** end-user feedback, deduplication
- **Data enrichment:** relationship inference (pub-project, citations, similarity, user behavior), content classification
- **Enhanced Pubs** (DANS, EBI, BADC pilots): compound info objects, import/export management, visualization

Services for research administrators





Services for Project coordinators



○ One click from the project's publications

○ Assist in progress reporting

– Create HTML, csv files to use in EC templates

These data are submitted, which means that they become available in publications and services from the beginning until the end of the project. Updates are possible at any time.

TEMPLATE A1: LIST OF SCIENTIFIC (PEER REVIEWED) PUBLICATIONS, STARTING WITH THE MOST IMPORTANT ONES

NO.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year of publication	Relevant pages	Permanent identifiers ¹³ (if available)	Is/Will open access ¹⁴ provided to this publication?
1	Economic Transformation in Hungary and Poland		European Economy	No 43, March 1990	Office for Official Publications of the European Communities	Luxembourg	1990	pp. 151 - 167		yes/no
2										
3										

- Use this HTML snippet to embed the project's **dynamic list** of publications in your site



```
<script type="text/javascript">
<!--
var oa_grant = '212488';
document.write('<div id="oa_widget"></div>');
document.write('<script type="text/javascript" src="/index.php?
option=com_openaire&view=widget&layout=bootstrap&format=raw&ga=' + oa_grant + '>
</script>');
-->
</script>
```

○ Provide code snippets to embed dynamic list of project publications in project site



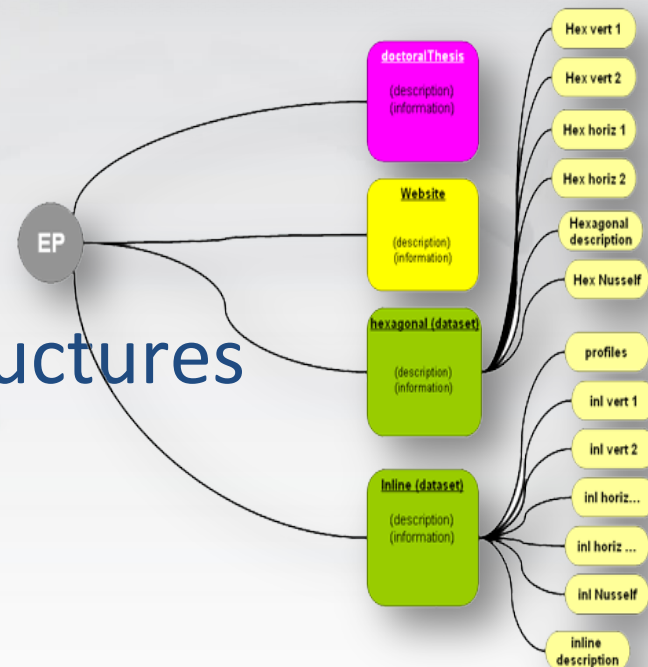
Services for data providers

Guidelines and tools

- Tools for data source validation and registration
- Guidelines for publication providers  
- Guidelines for data providers - accessing (metadata of) datasets from providers while minimizing effort to comply
 - *Metadata* (synergy with DataCite, EuroCRIS, EBI-EMBL, DANS, STFC/BADC): indications on **minimal metadata** about datasets (e.g., identifiers, date of creations, title, URLs) and best-practices for interlinking datasets and publications
 - *Access protocols*: no requirements for adopting precise protocols (e.g., OAI, FTP) or ID/URL frameworks (e.g., OpenURL, DOI) to comply

Services for researchers – “Enhanced” publications

- A publication in context
 - Semantic links to funding, data, publications, ...
- Cross-discipline model
 - EBI-EMBL, DANS, BADC
 - Export to other data infrastructures





Services for scientists

○ Subscribe to topics of interest

Add New Subscription

Select topic: **Projects**

- Projects
- Non-SC39 Projects
- SC39 Projects
- Publications in programme...
- Publications in scientific area...
- Claimed Publications
- Publications in project...
- Publications in non-SC39 projects
- Publications in SC39 projects
- Repositories in...

Send an alert to: available.

Alert using:

Post alert to: administrator interface.

○ Receive notifications through various means

Alert using: **Atom Feed**

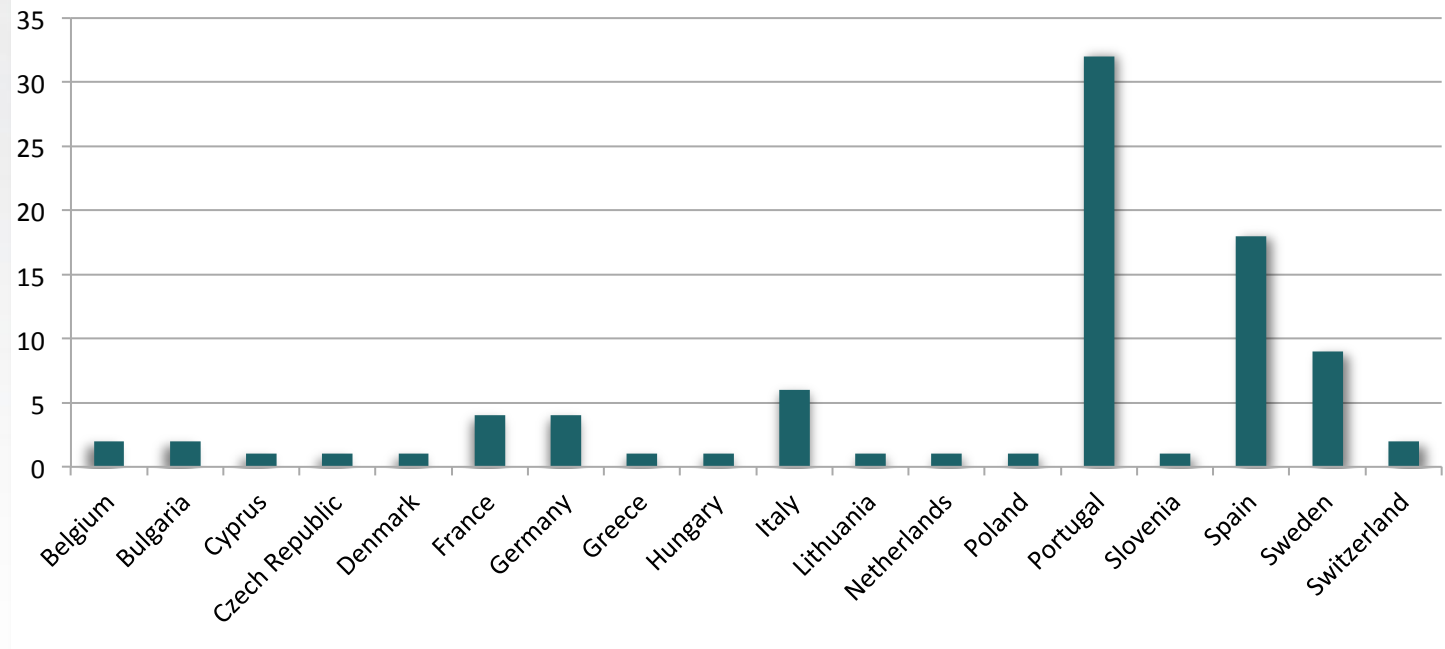
- Atom Feed
- My News
- OpenAIRE News
- Email
- twitter

Post alert to: specified in the administrator interface.



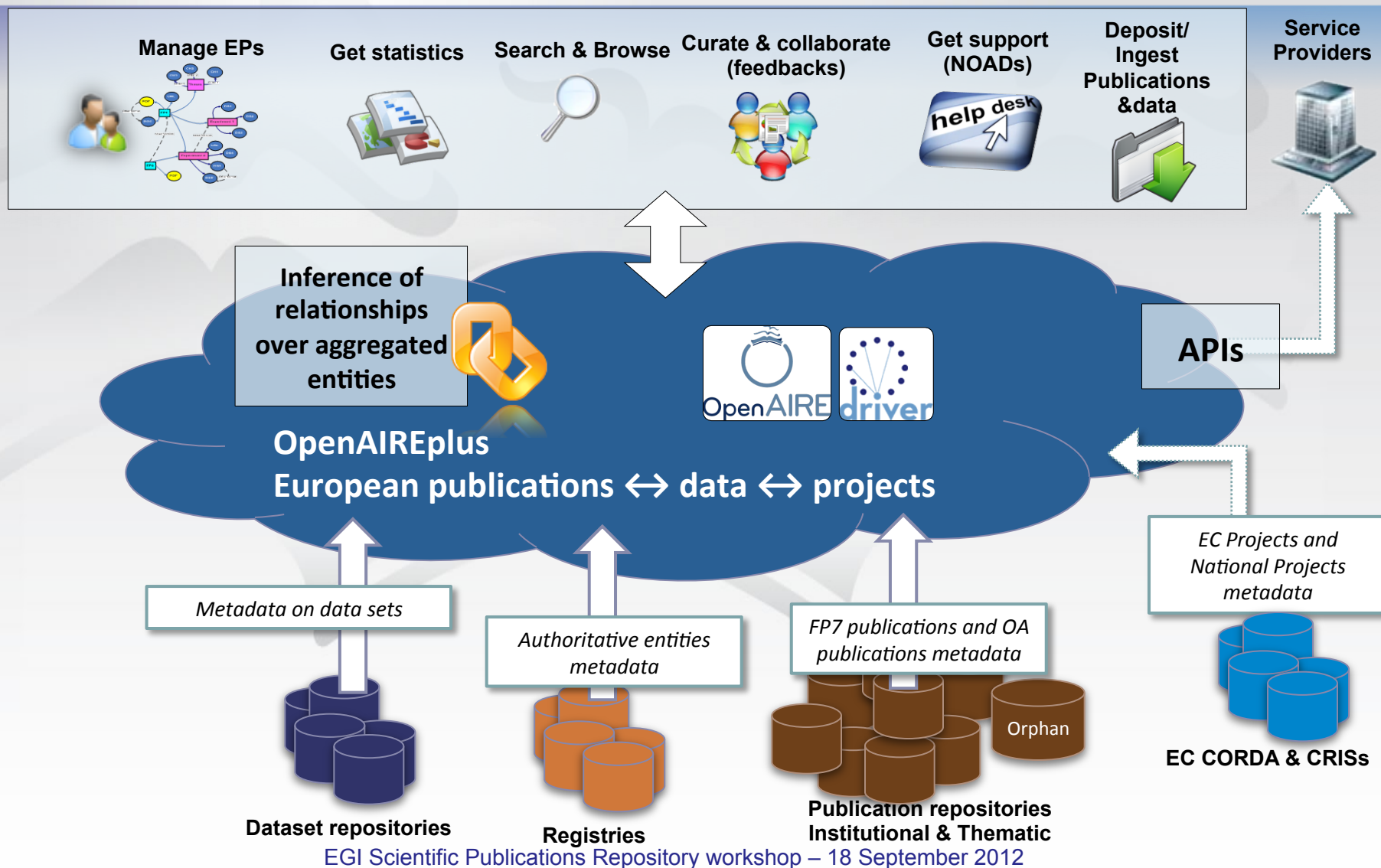
OpenAIRE in numbers

- > 26k FP7 publications from ~ 5000 projects
- > 10k Open Access
- 94 Open AIRE compliant OA Repositories & Journals





Production Infrastructure (24x7)





Background



Enabling Technology



D-NET Software toolkit



- **Service-oriented data infrastructure enabling technology**
- Adoption: Projects (DRIVER, OpenAIRE, EFG, HOPE, EFG1914) and nations (Spain-Recolecta, Poland, Belgium, Argentina (in progress))
- By: CNR-ISTI (IT), U Athens (GR) , ICM (PL), U Bielefeld (GE)



INVENIO Repository



- **Customizable repository platform:** workflows and data models
- Adoption: CERN digital library and 30+ institutions world-wide
- By: CERN (Switzerland), collaboration from DESY, EPFL, FNAL, SLAC



D-NET Software Toolkit

- Enabling data infrastructure software
 - **Registry:** resource (data sources and services) registration and discovery
 - **Authentication** and **authorization** of resources
 - **Workflow management:** autonomicity by orchestration and event handling
 - **Openness:** «easy» integration of new resource typologies
- Rich kit of data management and access services



Partners and Roles

○ Coordination

- University of Athens - GR
- Goettingen University Library - DE
- CNR-ISTI - IT

○ Technical production & operation

- 5 partners with expertise in technologies for Digital Libraries and Data Infrastructures

○ General

- Starting date: Dec 1, 2011
- Duration: 30 months
- Total budget: 5.2 Mi

○ Scientific communities

- EBI – biology
- DANS – social sciences
- STFC/BADC – climate

○ Networking Organization

- 5 libraries, active in OA movement

○ National Open Access Desks

- All member states
- Norway, Switzerland, Turkey, Iceland



Ongoing and under discussion Liaisons

○ DataCite

○ OAJ

○ e-IRG

○ CODA

○ REsearch.eu

○ CERIF

○ OpenDOAR

○

○ Mendeley

○ e-IRG

○ ORCID

○ EUDAT

○ RepositoryNET+

○ ...



Questions?

For more:

OpenAIRE infrastructure: <http://www.openaire.eu>

DRIVER infrastructure: <http://www.driver-community.eu>

D-NET Software toolkit: <http://www.d-net.research-infrastructures.eu>

INVENIO Repository: <http://invenio-software.org>