

Donatella Castelli

Consiglio Nazionale delle Ricerche
Istituto di Scienza e Tecnologie dell'Informazione

OpenAIRE: Open Access Infrastructure for Research in Europe

Open Access policies by HORIZON 2020 and
international issues regarding open science

October 13th-14th, 2015 – Kobe University, Japan





Outline

- **EC Open Access mandate**
- **OpenAIRE Infrastructure overview**
- **Future steps**



EC Open Access mandate



“The European Commission’s vision is that information already paid for by the public purse should not be paid for again each time it is accessed or used, and that it should benefit European companies and citizens to the full.”

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

2008: OA EU Pilot Mandate



Seventh Framework Programm OA Pilot

- Grant agreement Special Clause 39
- 20% programme areas
- Deposit in repositories an electronic copy of the published version or the final manuscript accepted for publication
- OA after embargo period (6-12 months)
- Gold payments during the project

OpenAIRE

Open Access Infrastructure for Research in Europe

An instrument to support and monitor the OA mandate

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

OpenAIRE Phase-1

OpenAIRE

Dec. 2009 - Nov. 2012

OpenAIREplus

Dec. 2011 – Dec. 2014

Infrastructure for:

- Promotion and support in the implementation of EC OA mandates for publications
- Identification, deposition, access of OA articles
- Linking of articles with data and other contextual information
- Monitoring of the level of OA compliance and production of indicators for decision makers



2014: Horizon2020 OA Mandate Progression

“Open access can be defined as the practice of providing on-line access to scientific information that is free of charge to the end-user and that is re-usable. In the context of research and innovation, 'scientific information' can refer to

- (i) **peer-reviewed scientific research articles** (published in scholarly journals) or
- (ii) **research data** (raw data, curated data and data underlying publications)”





EU funded projects Grant Agreement Art. 29.2 - OA to publications

Each beneficiary must ensure open access to all peer-reviewed scientific publications relating to its results

- **Deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication** in a repository for scientific publications
- **Ensure open access** to the deposited publication and its metadata— **(max embargo period 6 months)**
- Publications must **be at least readable online, downloadable and printable** - **Additional rights** should be provided, e.g. the right to copy, distribute, search, link, crawl, and mine
- Open access **to other types of scientific publications** (e.g. monographs, books, conference proceedings and grey literature) should be encouraged

(OpenAIRE is the recommended entry point for researchers to determine what repository to choose. It also offers support services for researchers, such as the National Open Access Desks.)

Art. 29.3 - OA to research data

H2020 Open Data Pilot

Beneficiaries participating in projects that have opted for the Data Pilot must:

- **Deposit in a research data repository** and take measures to make it possible for third parties to **access, mine, exploit, reproduce and disseminate, free of charge** for any user, the following:
 - the data, including associated metadata, needed to validate the results presented in scientific publications;
 - other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan'
- **Provide information, via the repository, about tools and instruments** at the disposal of the beneficiaries and necessary for validating the results (and, where possible, provide the tools and instruments themselves.)

(It is expected that OpenAIRE will become an entry point for linking publications to underlying research data)

OpenAIRE Phase-2

OpenAIRE2020
Jan. 2015 a Jun. 2018

OpenAIRE Legal entity

- **Extended infrastructure:**
 - Framework for the implementation of the Open Data Pilot for H2020
 - Pan-European OA Research Information Management System
- **New impact measurements metrics**
 - Scholarly performance multidimensional indicators
 - Social, cultural and economic impact indicators
- **Liaison with similar initiatives around the world**

OpenAIRE

OPEN ACCESS INFRASTRUCTURE



Human Network

www.openaire.eu



Technical Infrastructure

Participatory Approach

Building on existing assets



Human Network
(developers, users, collaborators)



ORGANIZATIONS

PEOPLE

REPOSITORIES

TECHNOLOGIES



Technical Infrastructure

Building and operating the infrastructure: a multi-disciplinary effort



National repository managers

- NOADs
- Guidelines
- Linking local to EU and global initiatives

Open Access experts

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

Data producers & users

- Linking to other data infrastructures
- Best practices for data discovery

Information & Computer Scientists

- Efficient e-Infra technologies development
- State of the art technologies (big data, linked data)

50 Project consortium partners

Legal experts

- Legal & policy recommendations

Serving many different roles

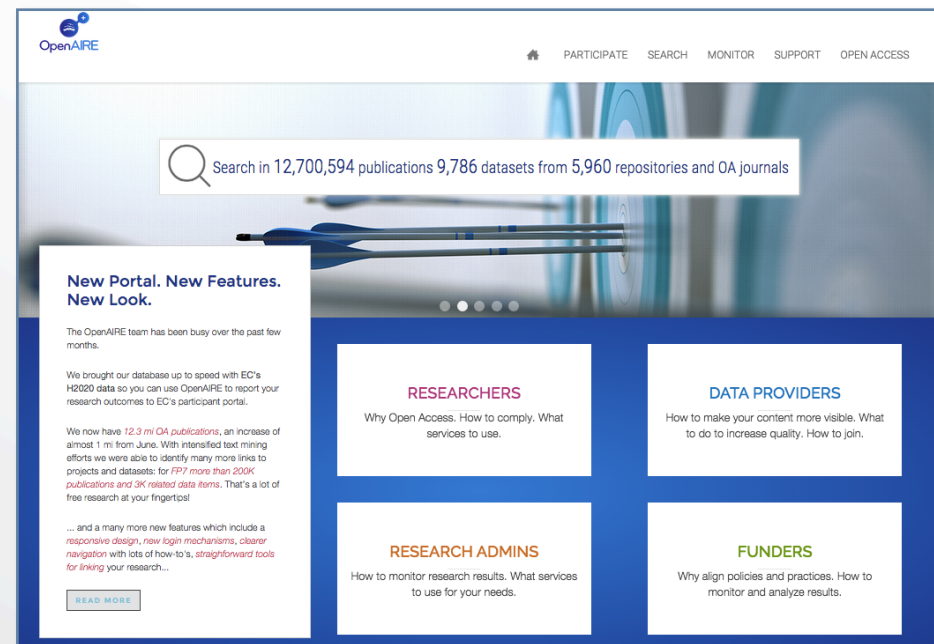
Research initiatives

Projects partners

Project coordinators

Data Center administrators

Research administrators



Repository managers

Policy Makers

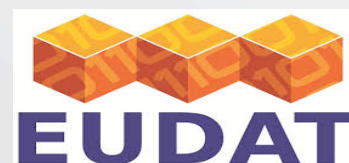
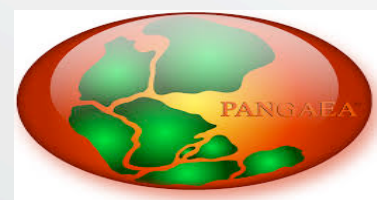
Funding agencies

Researchers

Publishers



Leveraging several collaborations



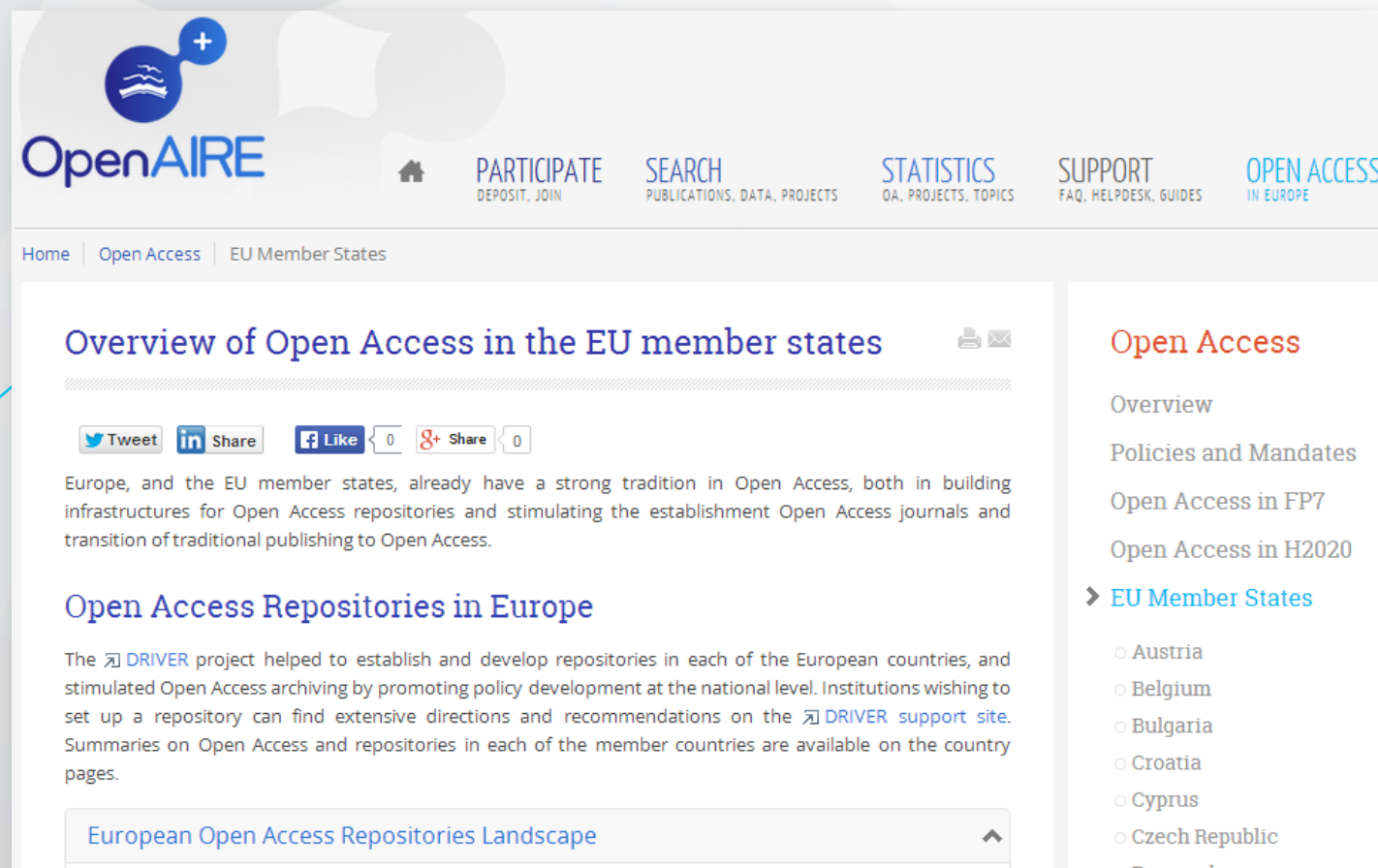
Participatory approach: different implementation models in EU Member States



- Strong national distributed repository networks
- Strong national repository network integrated with CRIS system
- Nationally centralized solutions
- Significant number of repositories and some “strong” ones

Flavours of Mandates

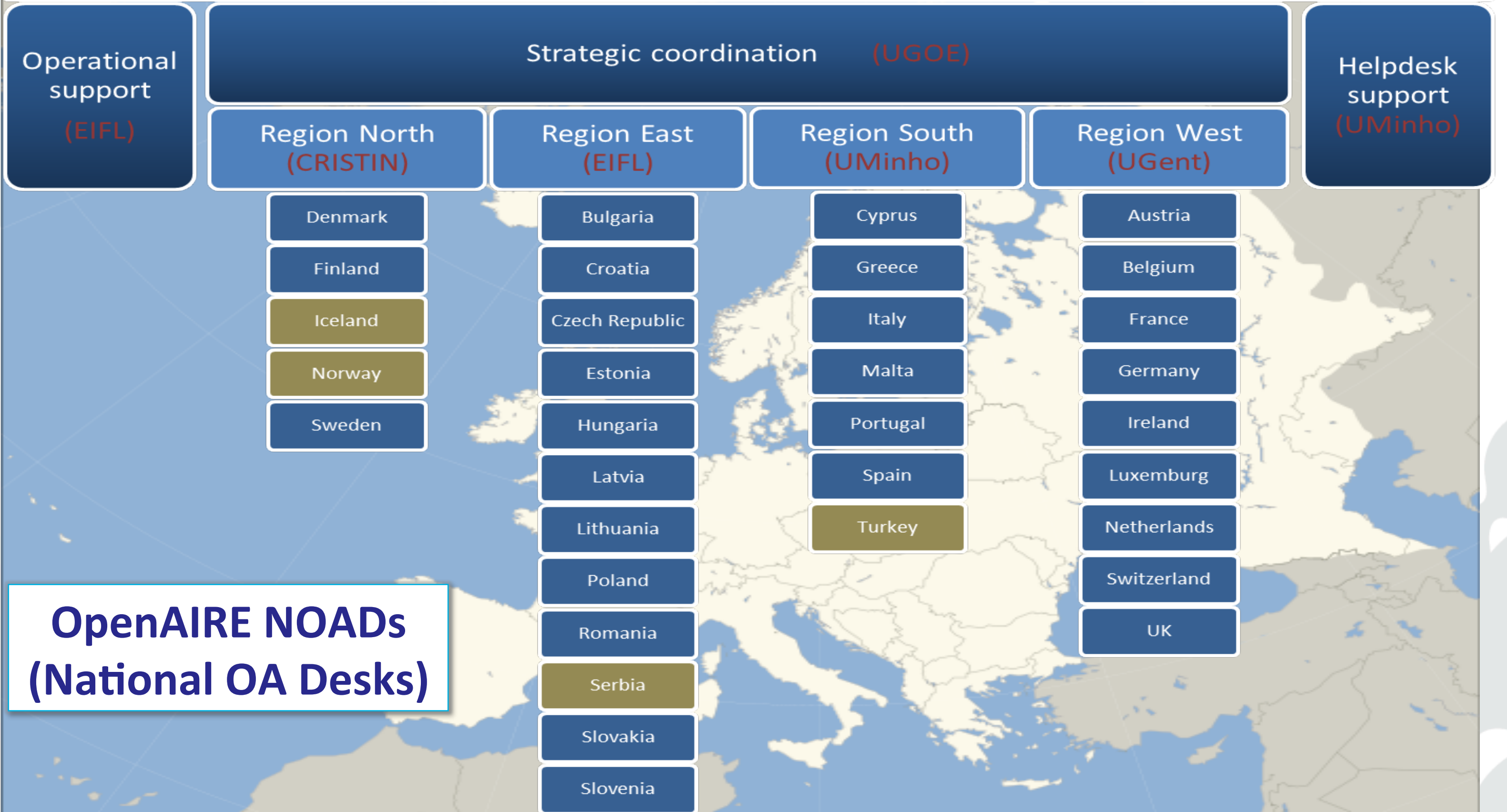
Diverse European landscape. Still in the ,Early Stages' of Development



The screenshot shows the OpenAIRE website interface. At the top, there is a navigation bar with the OpenAIRE logo and several menu items: Home, Open Access, EU Member States, PARTICIPATE (DEPOSIT, JOIN), SEARCH (PUBLICATIONS, DATA, PROJECTS), STATISTICS (OA, PROJECTS, TOPICS), SUPPORT (FAQ, HELPDESK, GUIDES), and OPEN ACCESS IN EUROPE. Below the navigation bar, the main content area is titled 'Overview of Open Access in the EU member states'. It includes social media sharing buttons for Twitter, LinkedIn, Facebook, and Google+. The text below the buttons states: 'Europe, and the EU member states, already have a strong tradition in Open Access, both in building infrastructures for Open Access repositories and stimulating the establishment Open Access journals and transition of traditional publishing to Open Access.' There is a section titled 'Open Access Repositories in Europe' which mentions the DRIVER project and provides a link to the DRIVER support site. On the right side of the page, there is a sidebar with a list of 'EU Member States' including Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, and Denmark.

- **Institutional mandates**
 - General agreement on benefits of gathering research output
 - Mandatory vs voluntary
- **National/funder mandates**
 - UK has partially adopted the Gold route
 - Many countries take the Green route
- **EC OA policy is a guiding framework**
- **Harmonisation of policies is progressing**

OpenAIRE2020 Networking Organisation



**OpenAIRE NOADs
(National OA Desks)**



NOADs contribution

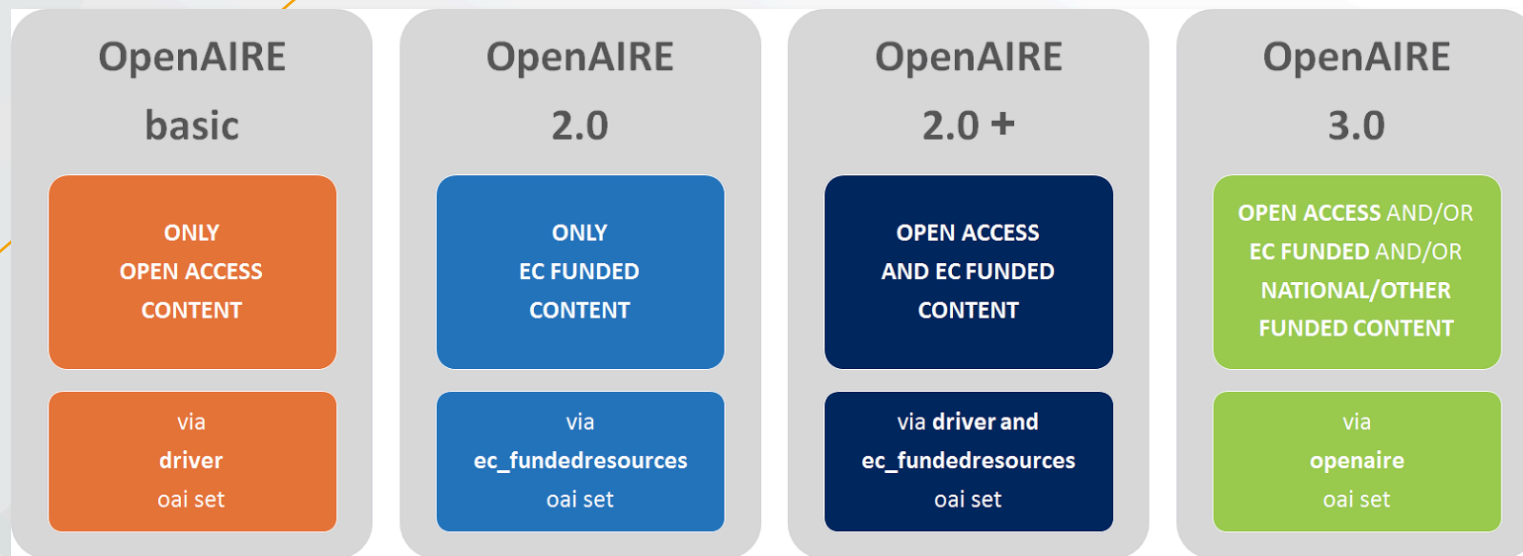
- **National coordination**
 - Research libraries and data centers
 - National research environments
- **OA advocacy through targeted activities**
 - Helpdesk, workshops, training
- **Harmonisation of OA policies**
- **Definition of shared guidelines**



OpenAIRE Guidelines



For data sources- what Repository are expected to provide



<http://guidelines.openaire.eu>

GUIDELINES FOR LITERATURE REPOSITORIES

based on Dublin Core

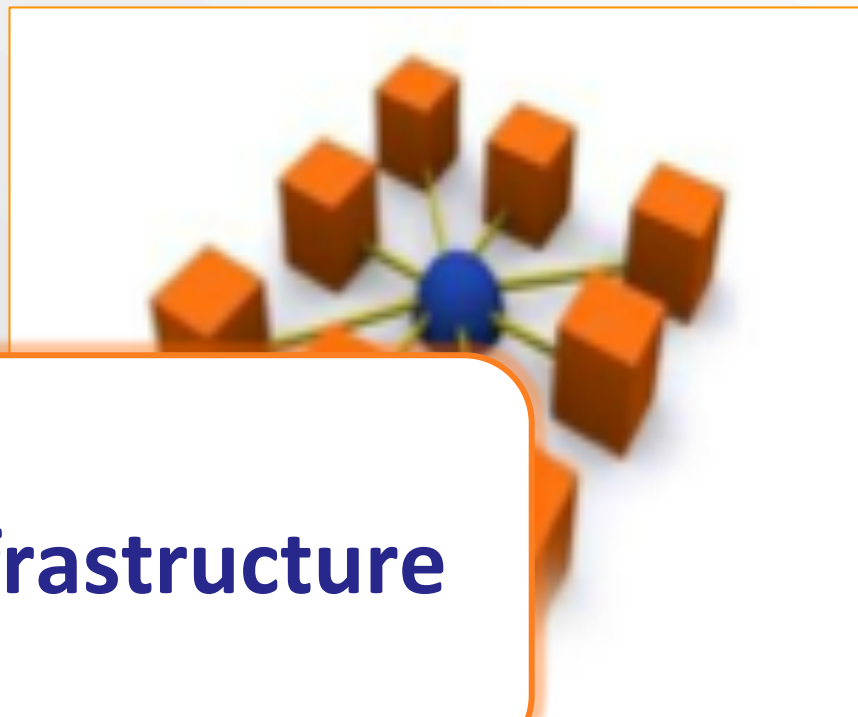
GUIDELINES FOR DATA REPOSITORIES

based on Datacite

GUIDELINES FOR CRIS MANAGERS

based on CERIF

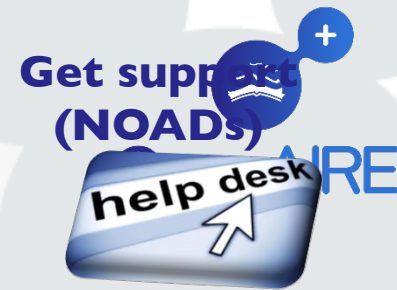
TECHNICAL INFRASTRUCTURE



e-Infrastructure



**Repository
service**



Claim/deposit Publications & data



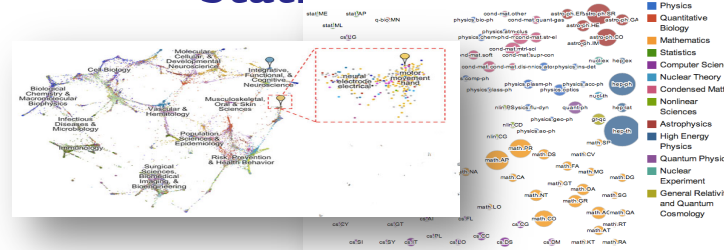
Search & Browse



Feedback

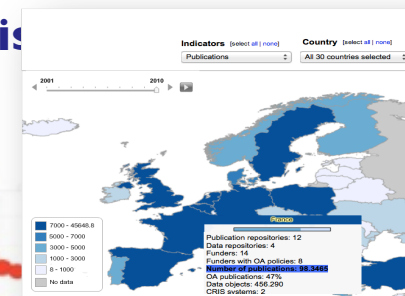
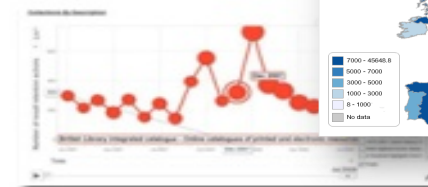


Linked Content Statistics



Research impact Citations, usage statistics

+++



Guidelines for use services

Validation



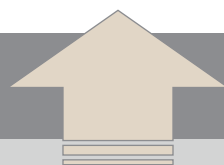
Cleaning & Transformation



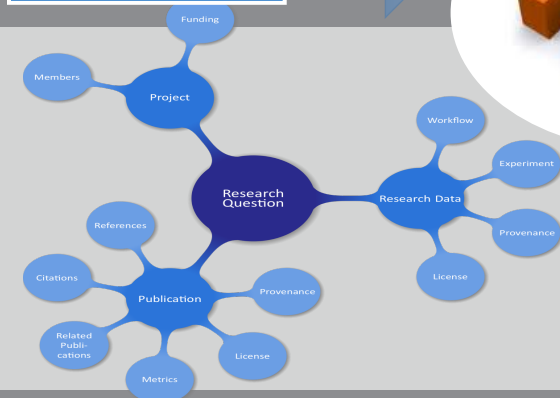
De-duplication



Metadata Enrichment



APIs



Guidelines for data interoperability

Metadata
And pdfs

Usage data



Publication repositories/ aggregators
Institutional & Thematic
Open Access Journals/Publishers

Institutional CRIS Systems

EC funding
National funding



zenodo

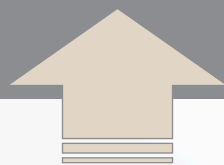
re3data
OpenDOAR

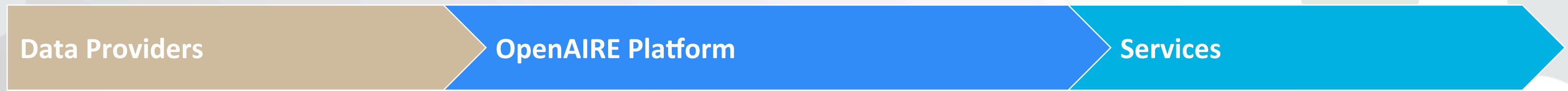
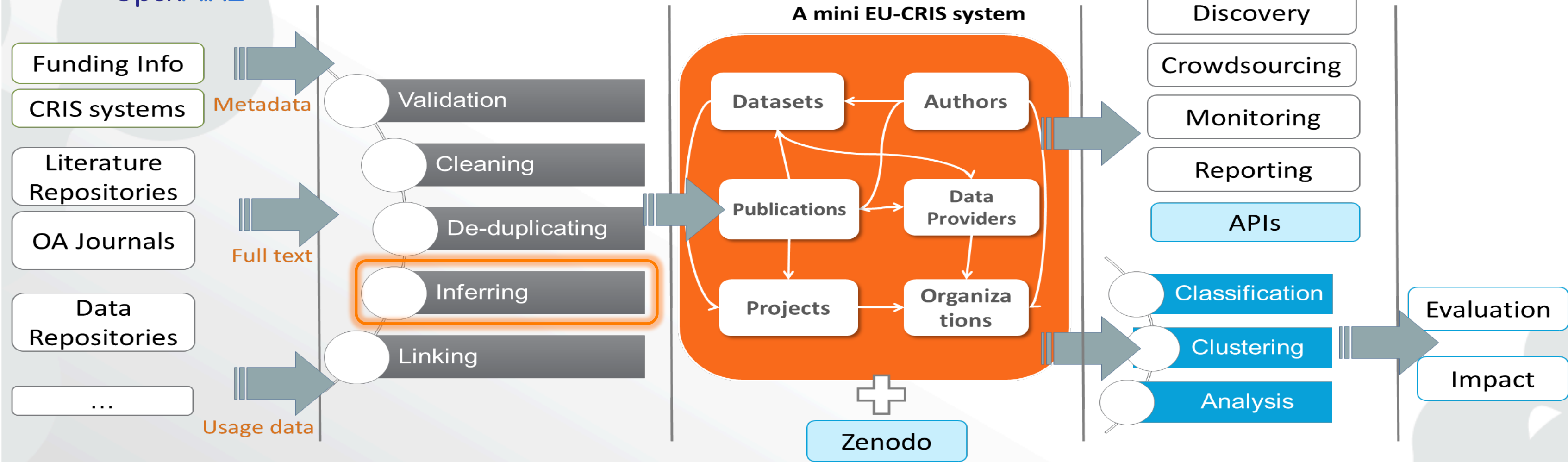
CERN/OpenAIRE "catch-all" repository

Metadata on data



Data repositories/ aggregators
Data Journals





Enrichment through mining

Examples

Property inference by mining publication PDFs

- Title of papers
- Authors of papers
- Affiliation of the authors at time of publication
- Subject classification



Relationship inference by mining publication PDFs

- Project-publication: EC and National projects (e.g. WT)
- Dataset-publication: DOI connections
- Publication-publication: citations by bibliography, text similarity



OpenAIRE information space

(September 2015)

- 12M publications (de-duplicated)
- 200,000 links publication-project from 5 funders
- 9,000 datasets linked to publications or projects
- 34,000 organizations (de-duplicated)



OpenAIRE information space: aggregated data sources (September 2015)



Datasource typology	Number of datasource	Type of objects
Journal Platform	5549	publications
Publication Repository	510	publications
Data Repository	38	datasets and publications
Aggregator of Publication Repositories	7	publications
Publication Catalogues	9	publications
Aggregator of Data Repositories	1	datasets
Aggregator/Publisher of Journals	4	publications
Entity Registry (data sources offering authoritative lists of entities)	11	data sources (publication repositories, data repositories), projects, funders
CRIS systems	0 (the first CRIS systems will be aggregated by the end of 2015)	publications, datasets, projects, persons



Services for a variety of users



HOME PARTICIPATE SEARCH MONITOR SUPPORT OPEN ACCESS

Search in 12,700,594 publications 9,786 datasets from 5,960 repositories and OA journals

New Portal. New Features. New Look.

The OpenAIRE team has been busy over the past few months.

We brought our database up to speed with EC's H2020 data so you can use OpenAIRE to report your research outcomes to EC's participant portal.

We now have *12.3 mi OA publications*, an increase of almost 1 mi from June. With intensified text mining efforts we were able to identify many more links to projects and datasets: for *FP7 more than 200K publications and 3K related data items*. That's a lot of free research at your fingertips!

... and a many more new features which include a *responsive design, new login mechanisms, clearer navigation* with lots of how-to's, *straightforward tools for linking* your research...

[READ MORE](#)

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.

RESEARCH ADMINS

How to monitor research results. What services to use for your needs.

FUNDERS

Why align policies and practices. How to monitor and analyze results.



Serving data providers & researchers

DATA PROVIDERS

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

- . Making repository content part of a EU information space
- . Check of compliancy to OpenAIRE Guidelines

RESEARCHERS

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

- . Deposit once and become compliant to the EC OA mandates
- . Be included in a European CRIS system
- . Automatic linking of own research outcome with contextual information.

Serving researcher administrators

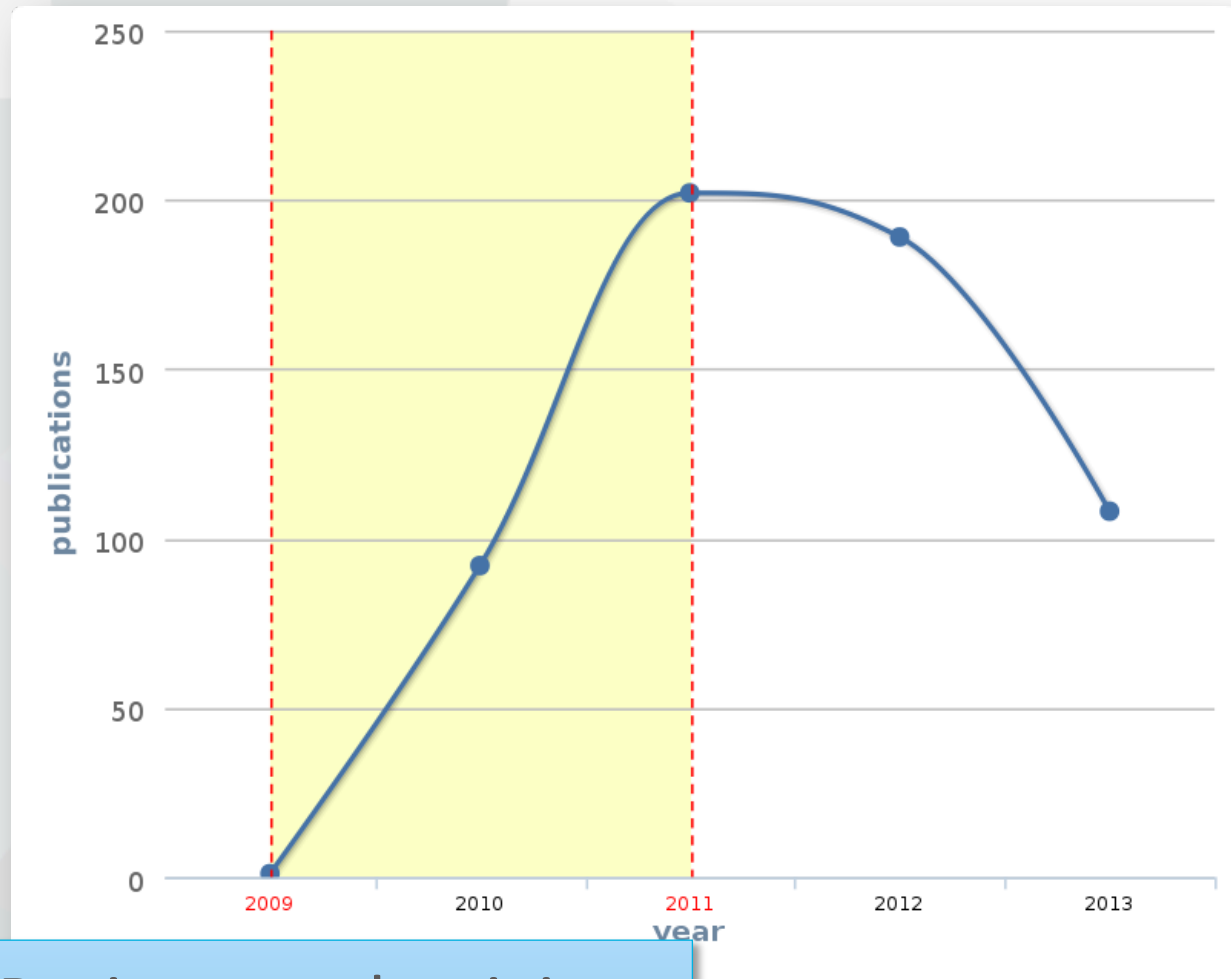
RESEARCH ADMINS

How to monitor research results. What services to use for your needs.

- SYGMA, the EC report system, can pre-fetch publications linked to projects
- EC project reports can be dynamically filled from OpenAIRE website
- Project web sites can use OpenAIRE widgets to dynamically show the list of related publications
- Project results can be compared with those of other projects

Serving funders: trends & indicators

FUNDERS
 Why align policies and practices. How to monitor and analyze results.

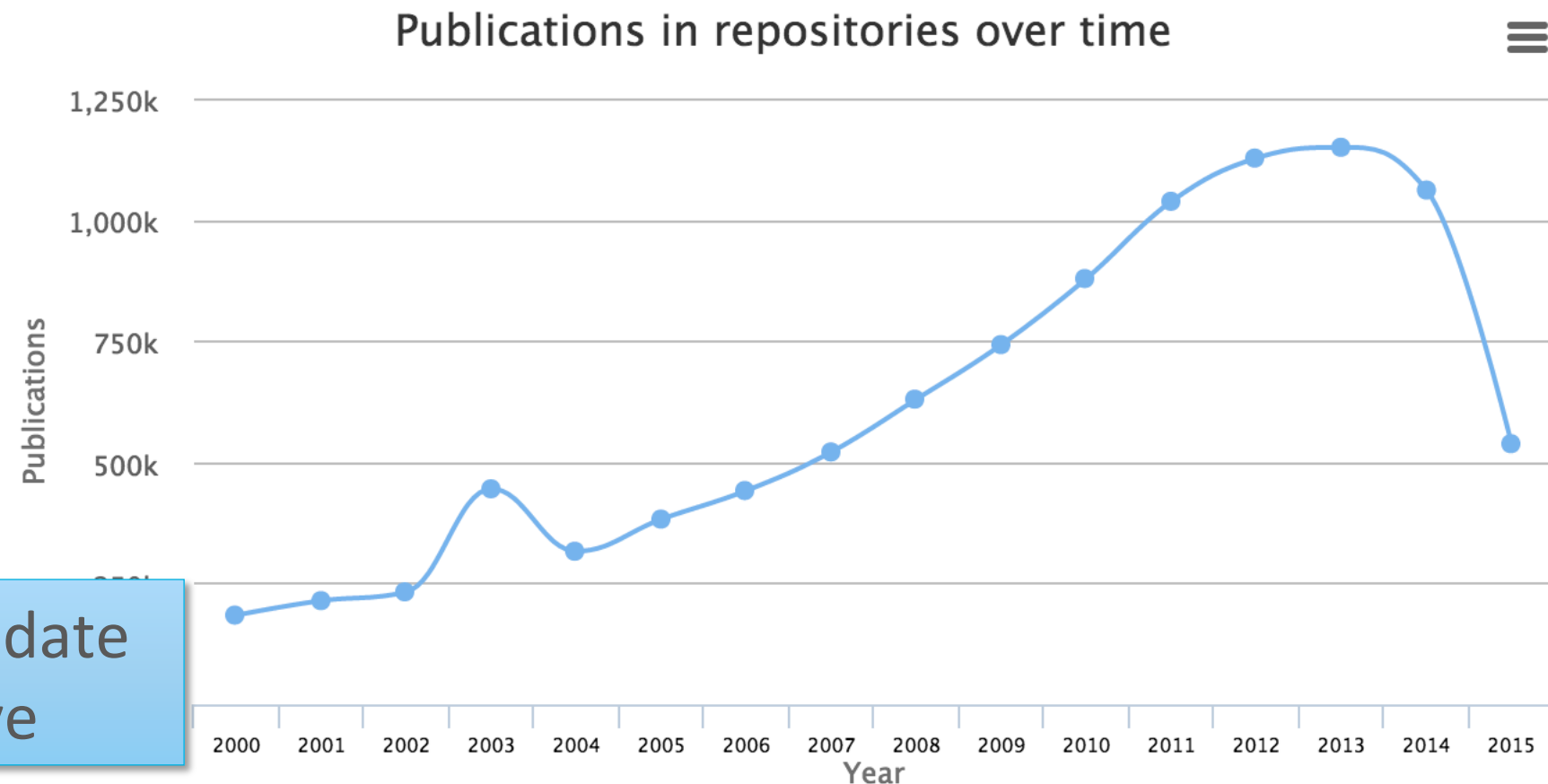


OpenAIRE Monitoring



12,700,594 publications from 658 data sources

12,295,576 published in OA and 234,502 of them are linked to funding

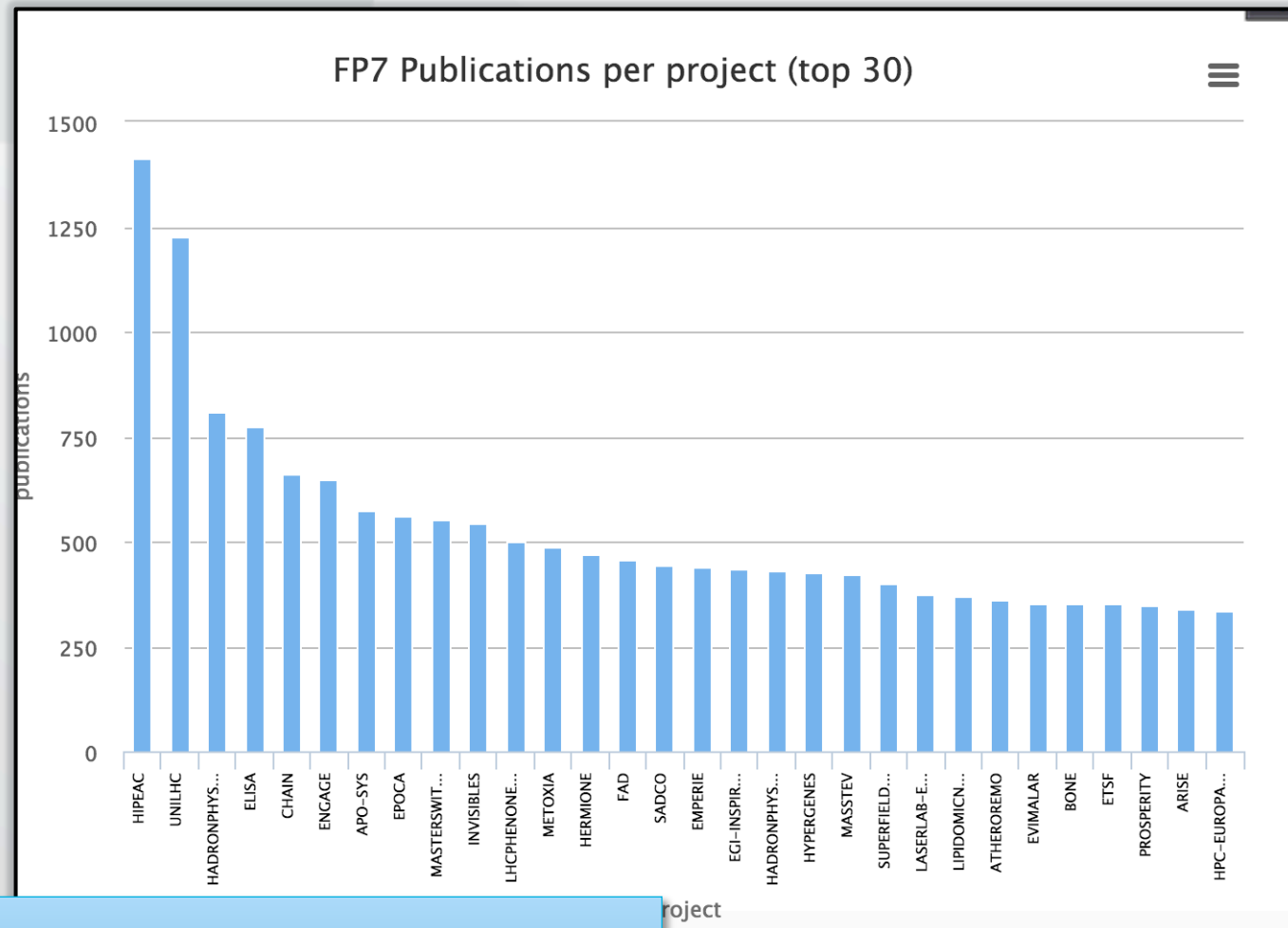


Project productivity over time

Post project-end monitoring

OA mandate curve





Comparing projects

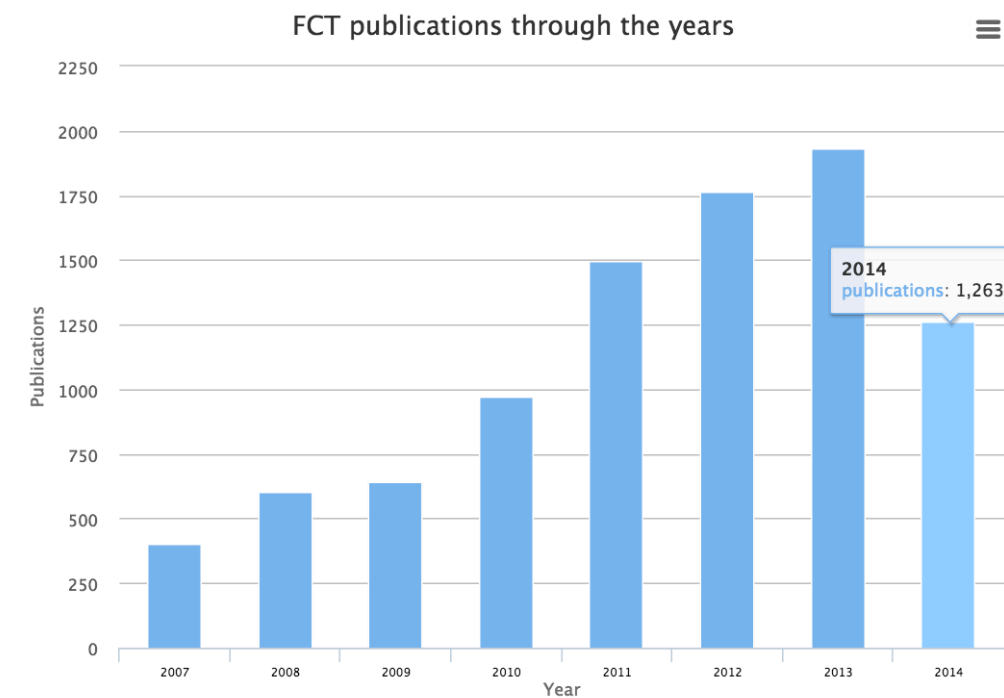
FCT - Fundação para a Ciência e a Tecnologia



10,087 publications in 3,575 projects (total of 10,688 projects)

9,017 are OA, 3 are restricted and 0 are still in embargo [Graph](#)

National Funders statistics





D-Net Enabling system



www.d-net.research-infrastructures.eu

- **Open Source (Apache License)**
- **Exploited by Argentina, Spain (La Recolecta) to create National aggregators**
- **Exploited also for other kind of infrastructures:**
 - **European Film Gateway (<http://www.europeanfilmgateway.eu/>),**
 - **HOPE – Heritage of Peoples’ Europe (<http://www.peoplesheritage.eu>)**
 - **EAGLE – Europeana Network of Ancient Greek and Latin Epigraphy (<http://www.eagle-network.eu>)**



Zenodo: the OpenAIRE offered repository

Publication
Poster
Presentation
Dataset
Image
Video/Audio
Software

Public Inielnsen-cern / decouple
forked from svenkreiss/decouple
33 commits 3 branches 3 releases 1 contributor
branch: master decouple /
ZENODO/GitHub integration demo
Inielnsen-cern authored 22 minutes ago latest commit 87d047f16c
Decouple Pass the common parameters explicitly to Leff scan and throw error i... a month ago
ModelGenerators Minor. 2 months ago
Plot Move Plot.utils to Decouple.src.plot_utils to make it available to ot... a month ago
output Init public repo. 2 months ago
plots Init public repo. 2 months ago
plotsForPaper Finer scan of robustness. Larger font size for eta arrow plots. a month ago
.gitignore Remove local LHCXSHiggsCouplings submodule and replace with depende... a month ago
.zenodo.json ZENODO/GitHub integration demo 12 minutes ago
LICENSE First version to work with pip. a month ago
Makefile Pass the common parameters explicitly to Leff scan and throw error i... a month ago
README.md ZENODO/GitHub integration demo 12 minutes ago
setup.py Bump version a month ago
Decouple and Recouple
DOI: 10.5281/zenodo.8345

Publishing sw:
GitHub-bridge

DOI 10.5281/zenodo.8345

```
{  
  "name": "Plehn, Tilman",  
  "affiliation": "Institut für Theoretische Ph  
  },  
  "description": "This repository contains the soft  
  "access_right": "open",  
  "license": "mit-license",  
  "related_identifiers": [{  
    "identifier": "arXiv:1401.0080",  
    "relation": "isCitedBy"  
  }]  
}
```

zenodo Research. Shared.
Search Deposit Browse Get started Help
lars.holm.nielsen@cern.ch
Home / Account / GitHub
Settings
Profile
Applications
GitHub
GitHub Repositories (updated 16 hours ago) Sync...
Get started
1 Flip the switch
2 Create a release
3 Get the badge
Select the repository you want to preserve, and toggle the switch below to turn on automatic preservation of your software.
Go to GitHub and create a release. ZENODO will automatically download a .zip package of all new releases and register a DOIs for them.
After your first release, you can get a DOI badge to include in your GitHub README file.
More question? Check out the FAQ.
Inielnsen-cern/altantis-conf OFF
Inielnsen-cern/dictdiffer OFF
Inielnsen-cern/decouple ON
Inielnsen-cern/flask-bower-grunt OFF
Inielnsen-cern/flask-cache OFF

Zenodo (cont.)



www.zenodo.org

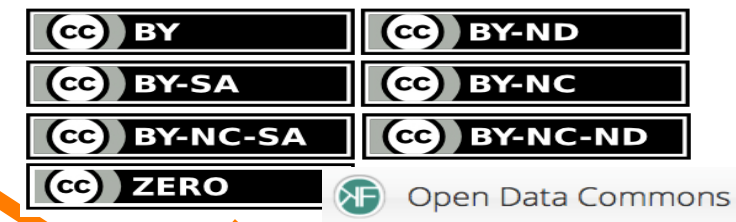
The screenshot shows the Zenodo interface for the article "Data Journals: A Survey" by Candela, Leonardo; Castelli, Donatella; Manghi, Paolo; Tani, Alice. The article is dated 30 January 2015 and is categorized as a "Journal article" and "Open access". The abstract discusses the challenges of data publishing and the role of data journals. The page includes a DOI (10.1002/asi.23358), keywords, publication date, and a list of grants including OPENAIREPLUS and IMAP. The article is published in the "Journal of the Association for Information Science and Technology" (2015). The page also features a "Files" section at the bottom with a table of file names and dates.



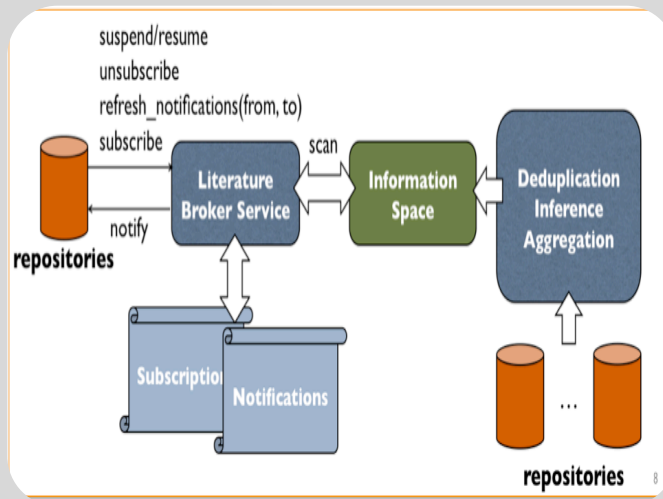
DOI 10.5281/zenodo.10594



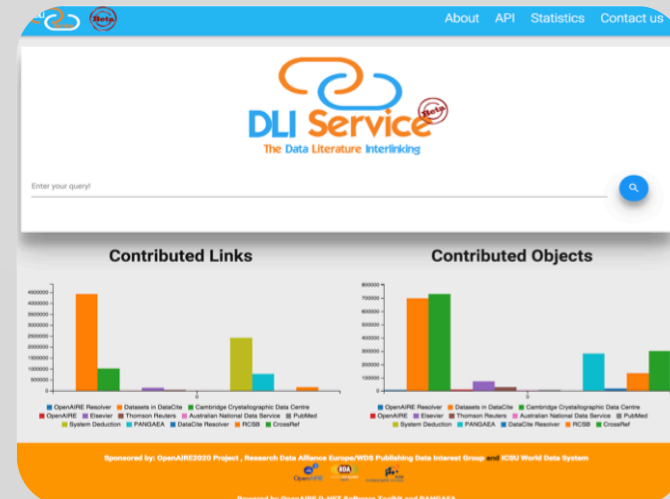
OpenAIRE-compliant metadata



A taste of on-going activities

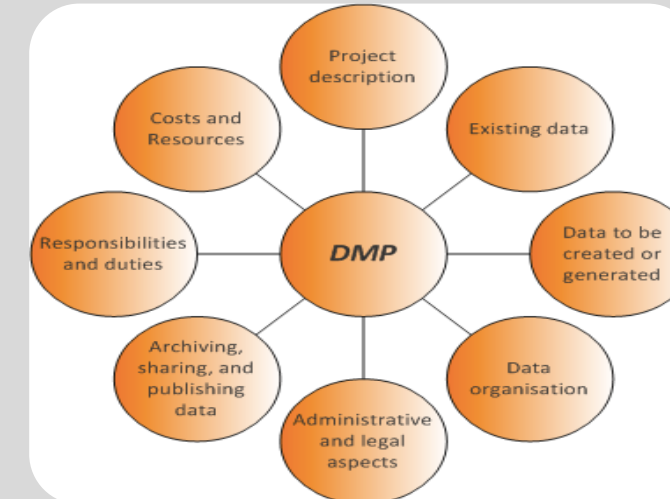


Literature
Broker
Service for
Institutional
Repositories



Data
Literature
Interlinking

<http://dliservice.research-infrastructures.eu>



Support to
Data Mgmt
Plan

A taste of on-going activities (cont.)



RDA-
ANDS
(Australia)



SHARE
(United
States)



La
Referencia
(South
America)

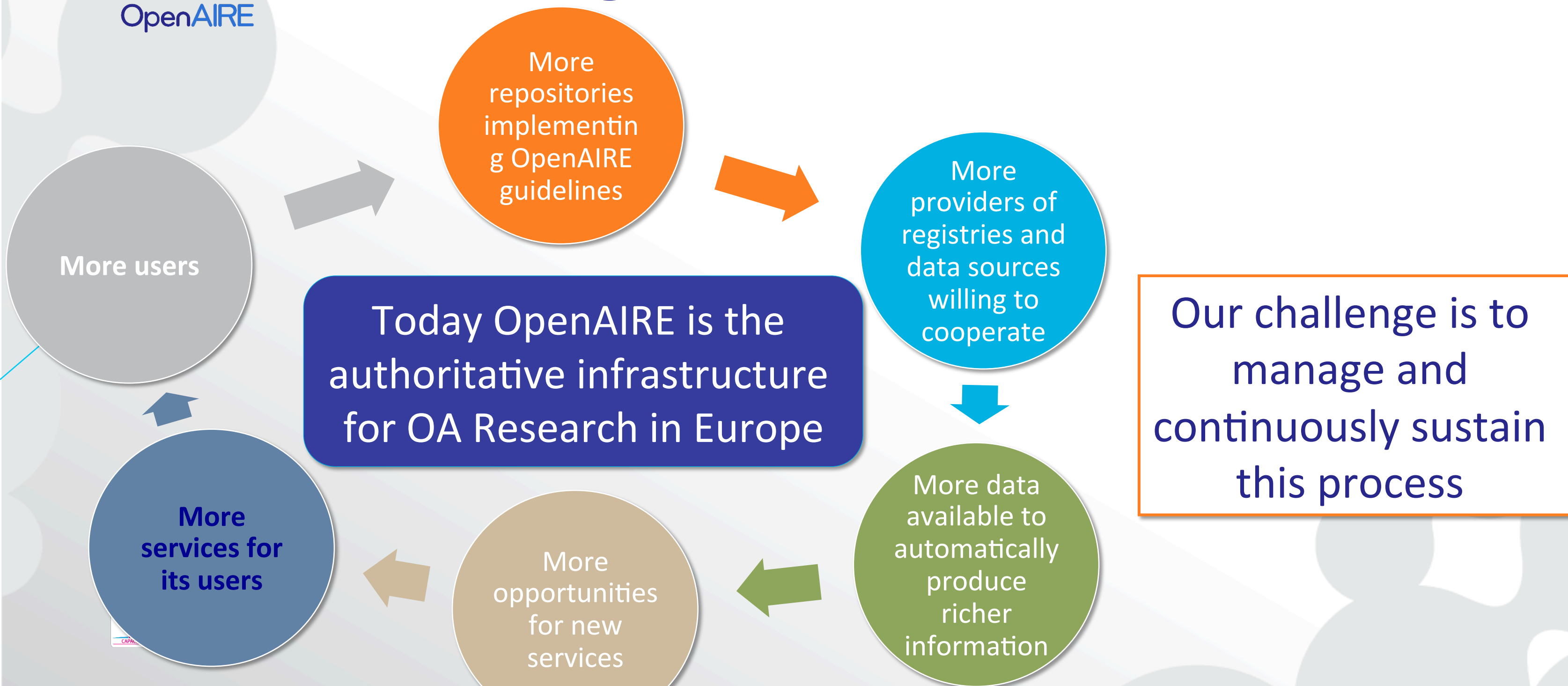


CAS
(China)


Global alignment of
repository networks
through COAR

- **Global voice for the repository community**
- **Facilitating interoperability across networks and repositories (and with other types of systems)**
- **Building capacity at the local level by developing and adopting value added services**

Concluding remarks



Thank you!

 www.openaire.eu

 @openaire_eu

 facebook.com/groups/openaire

 linkedin.com/groups/OpenAIRE-3893548

 donatella.castelli@isti.cnr.it