# OpenAIRE

Technical services & Dashboards for Institutions and Funders

Paolo Manghi

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

## e-infrastructure services



#### **Content providers**



#### Researchers



Third-party services (SMEs)



#### **End-User Services**



#### Info Space Services

Cleaning

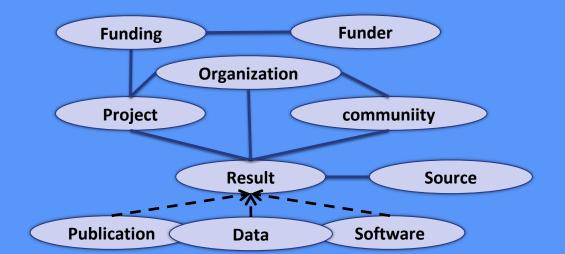
**Validation** 

**Harvesting** 

**De-duplication** 

Inference

Uploading





#### **Content Providers**

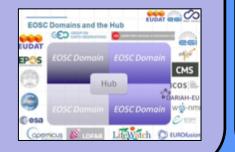
Publications Data repositories

Software OA repositories Journals

Registries



#### **EOSC-Hub**



#### Research Infras







# Catch-All Broker Service (CAB)

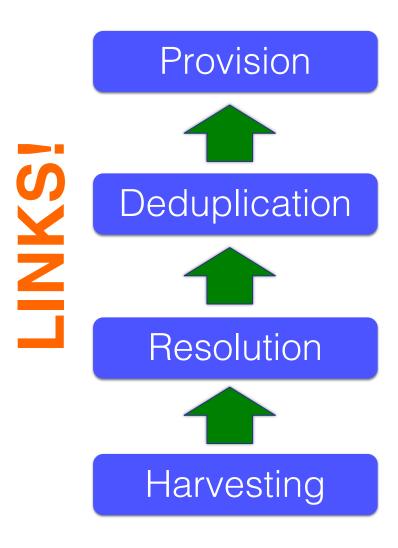


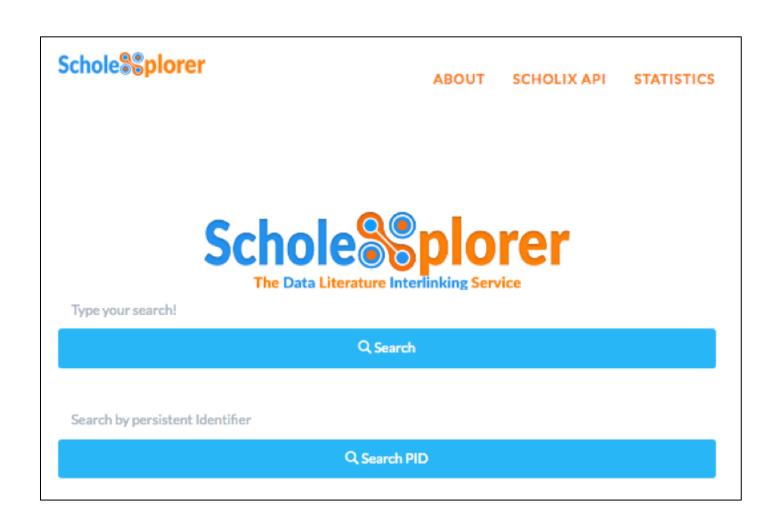
- Additions: New metadata records they were unaware of
- Enrichments: Extra information about records they already have
- Alerts: Wrong information in the records they already have





# Scholexplorer: data-literature links aggregator







Deployment cofunded by RDA

Search and browse in 880.000 literature objects, 5.840.000 datasets, 31.000.0000 bi-directional scholix links,

from 13000 publishers, 10 data centers, CrossRef, DataCite, and OpenAIRE.





http://scholexplorer.openaire.eu

## Scholix: use-case

## Scopus®





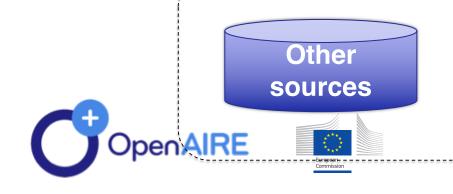


#### Link search/resolution







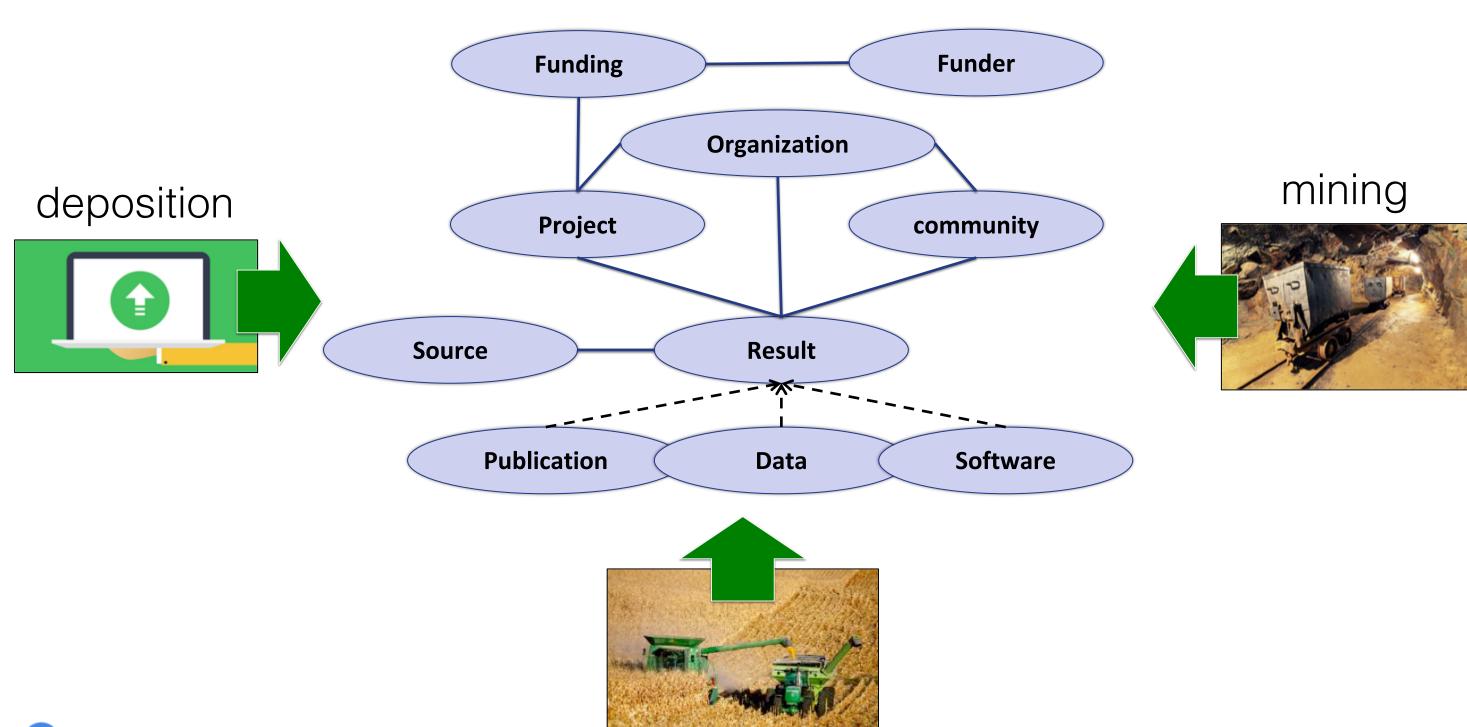








# **OpenAIRE Data Model and Flows**







# Zenodo: March 2018 report

### Content (205,000 records)

Datasets: 23.000

Software: 17,000

Publications: 137,000

Images: 178,000

Other: 7,000

#### Infrastructure

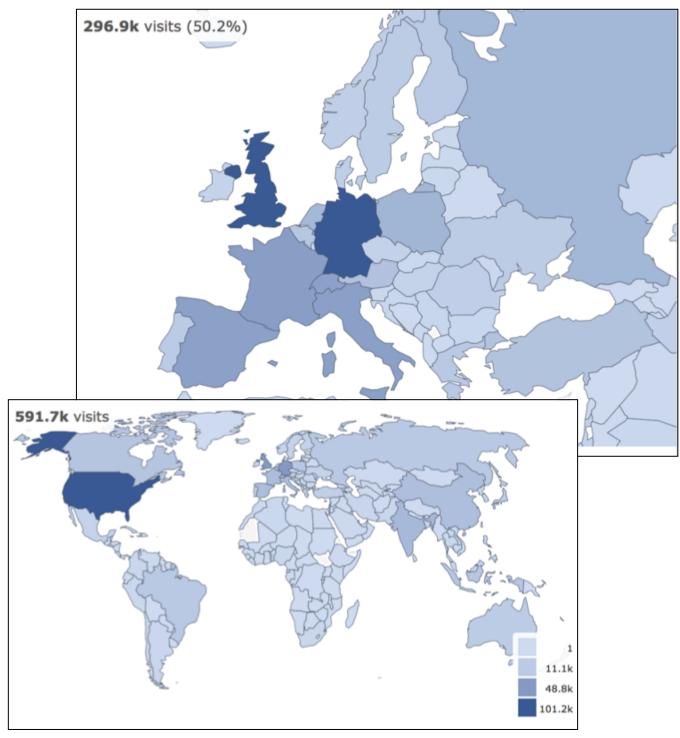
• 30 nodes, 105 TB

#### Usage:

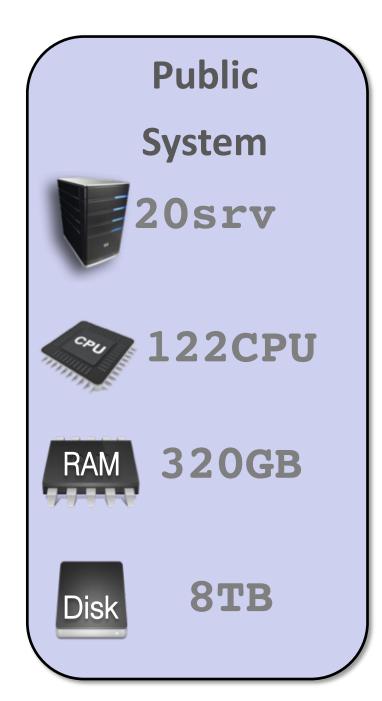
592k visits in 2017

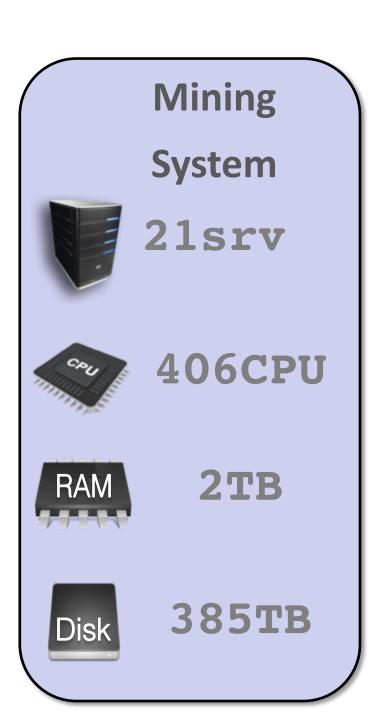


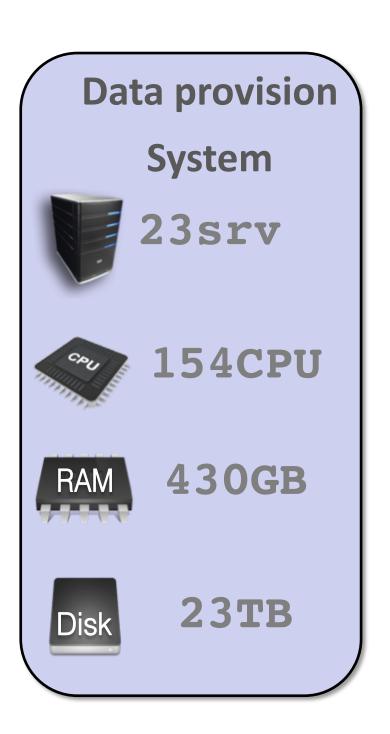




## e-infrastructure numbers







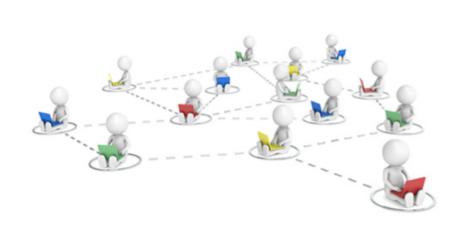


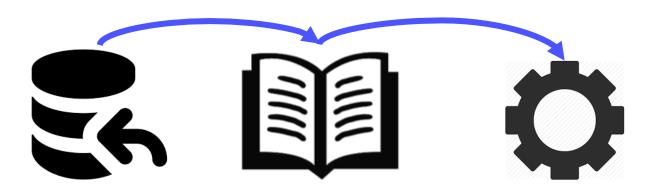




# Moving towards Researchers Research Community Dashboard

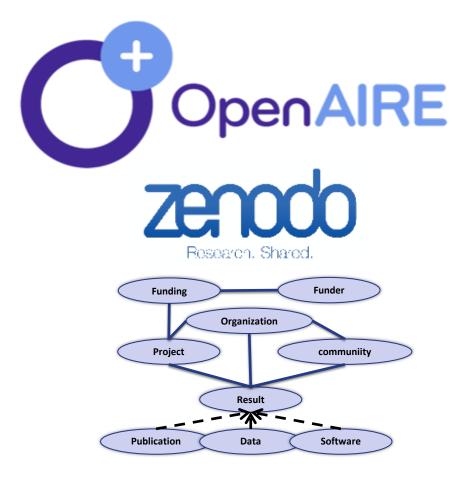
Data Publications Software







**Accessing via RCD and APIs** 





#### Research communities

#### On board

European Marine Sciences (Pangaea and Atlas community)

Cultural Heritage and Digital Humanities (the PARTHENOS research infrastructure)

Neuroinformatics (France Life Imaging national infrastructure)

Fisheries and aquaculture management (the BlueBridge and MARBEC infrastructures)

Environment & Economy (national/EU node of the United Nations Sustainable Development Solutions Network) EGI (Marios Chatziangelou, Giuseppe Della Rocca)

INSTRUCT (Evangelia D. Chrysina, Antonio Rosato, Chris Morris)

**RDA (Daniel Bangert)** 

**AgInfra+ (Panagiotis Zervas)** 

**CLARIN (Maria Eskevich)** 

#### **Bootstrapping**

ELIXIR (Rafael Jimenez, Theodore Dalamagas, Thanasis Vergoulis)

OpenRiskNet (Haralambos Sarimveis, Thomas Exner)

Ecology and Evolution (Antica Culina, KNAW)

EGA (Jordi Rambla, Mauricio Moldes)

**DARIAH (Marco Raciti)** 

SoBigData.eu (Fosca Giannotti)

**EPOS (Michele Manunta)** 

RISIS (Philippe Laredo)





# Aggregation as-a-service





# Broker as-a-service





# **OpenAIRE: D-NET Software Toolkit**

Technology transfer to third-party organizations





- National repository aggregation systems
  - Spain, La Recolecta
  - Poland, CEON
  - Turkey
  - Argentina
  - Ireland (ongoing)
- Thematic aggregator
  - World Data System aggregator (ongoing work)
  - EFG, HOPE, ESPAS, EAGLE





# OpenAIRE technical staff (40+ members)





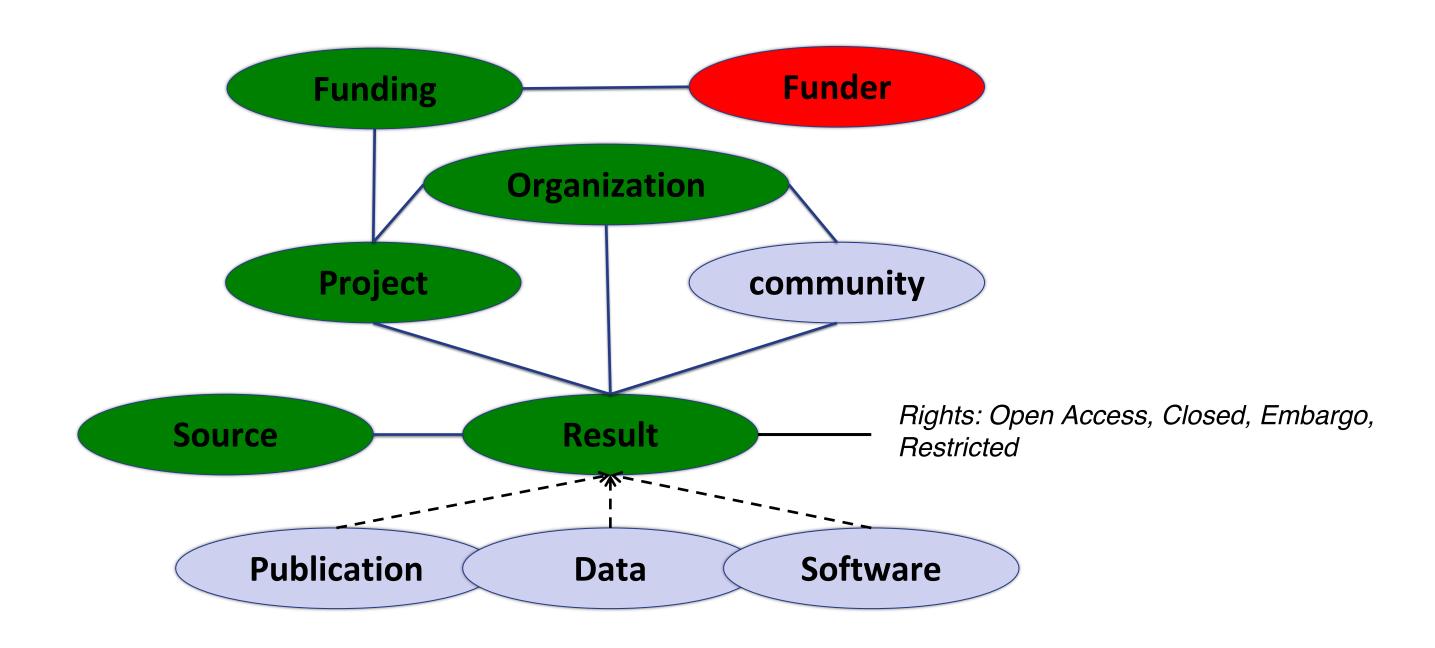


# Funder Dashboard





# Funders and OpenAIRE Data Model







# **Aims**

# Funding Impact

Publications, research data, software published thanks to grants awarded by the funder

# Open Access/Science Impact

Monitoring of Open Science impact: data/software FAIRness, reproducibility trends





## **Added-value functionalities**

#### Funders

Trends in research fields: new (multidisciplinary) disciplines

#### Institutions

OA/OS behavior, ability to attract cross-funder grants

## Projects

Success, interconnections, possible liaisons





## Users

#### Generic users

- Search & Browse
- Visualize Stats
- Claim (registered users)

# Managers (registered/authorized users)

- Alerts (e.g. violation of OA mandates)
- Impact propagation





# Stats for publications

# Funders' page for publications

Openness

**Publications Distribution** 

Publications by Scientific area

Publications per Project

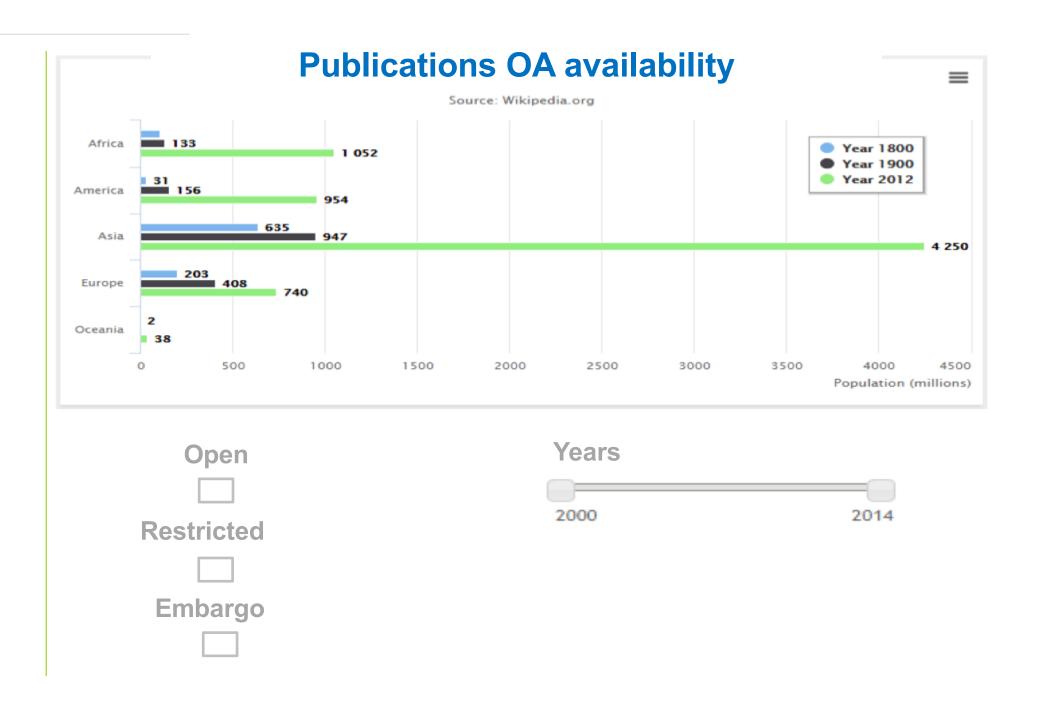
Publications by Document Type

Publications by Language



#### **Openness**

Publications Distribution
Publications by Scientific area
Publications per Project
Publications by Document Type
Publications by Language

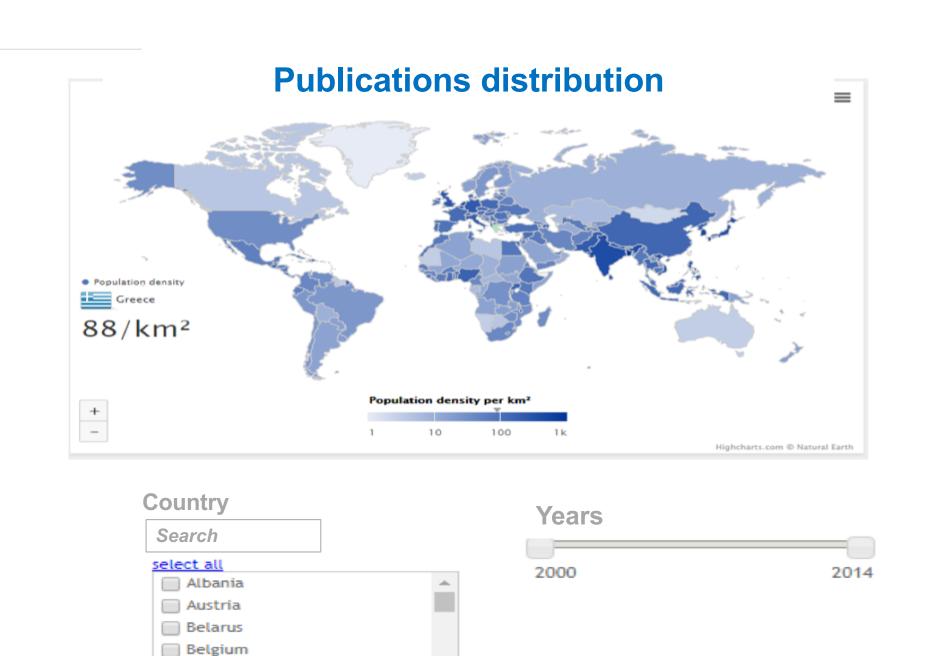




#### Openness

#### **Publications Distribution**

Publications by Scientific area
Publications per Project
Publications by Document Type
Publications by Language



Bosnia and Herzegovina

BulgariaCroatia



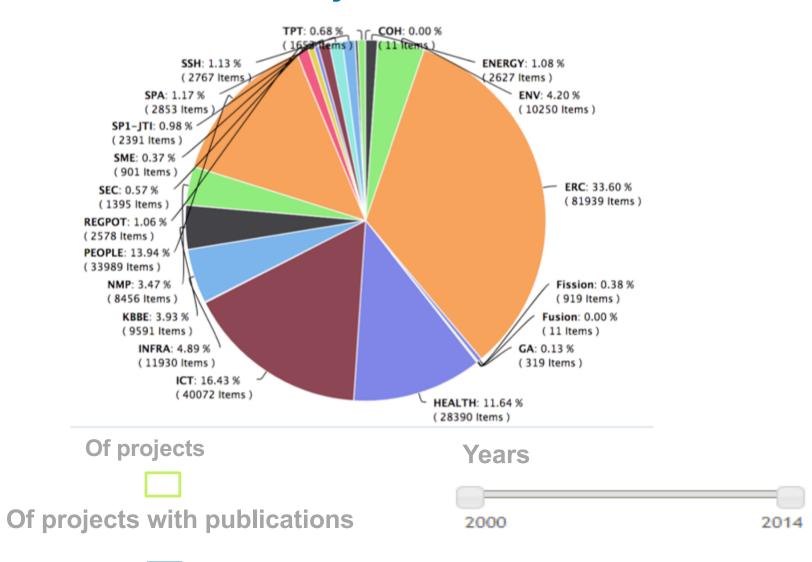
#### Openness

**Publications Distribution** 

#### **Publications by Scientific area**

Publications per Project
Publications by Document Type
Publications by Language

#### **Publications by scientific area**



Of projects with publications beyond their life



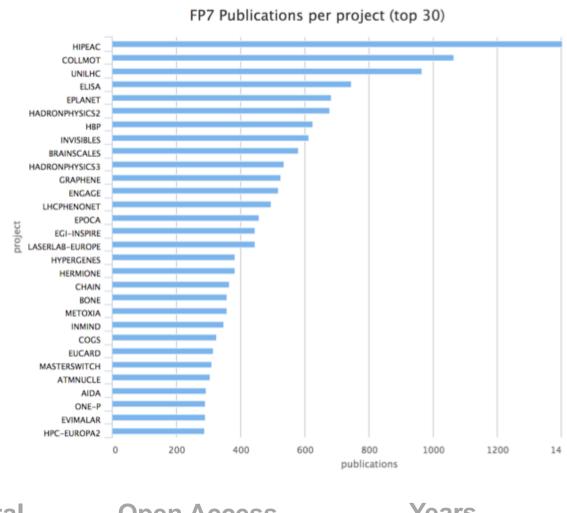
#### Openness

Publications Distribution
Publications by Scientific area

#### **Publications per Project**

Publications by Document Type Publications by Language

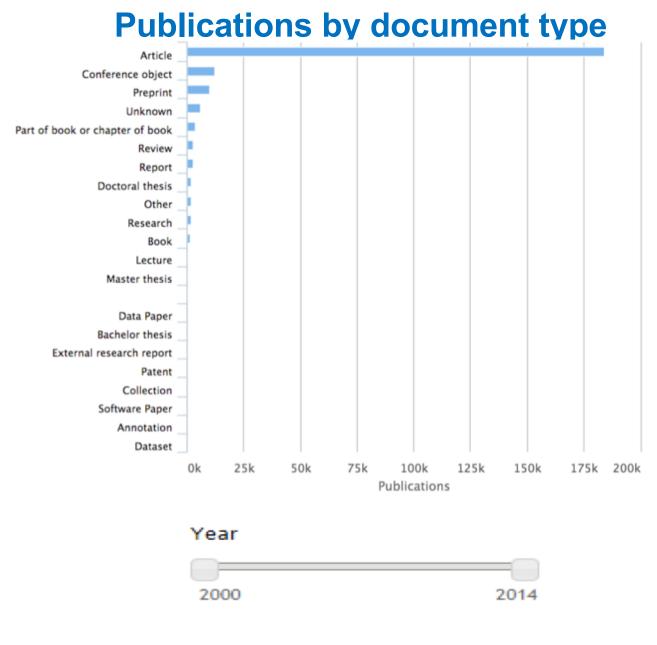
#### **Publications per project**







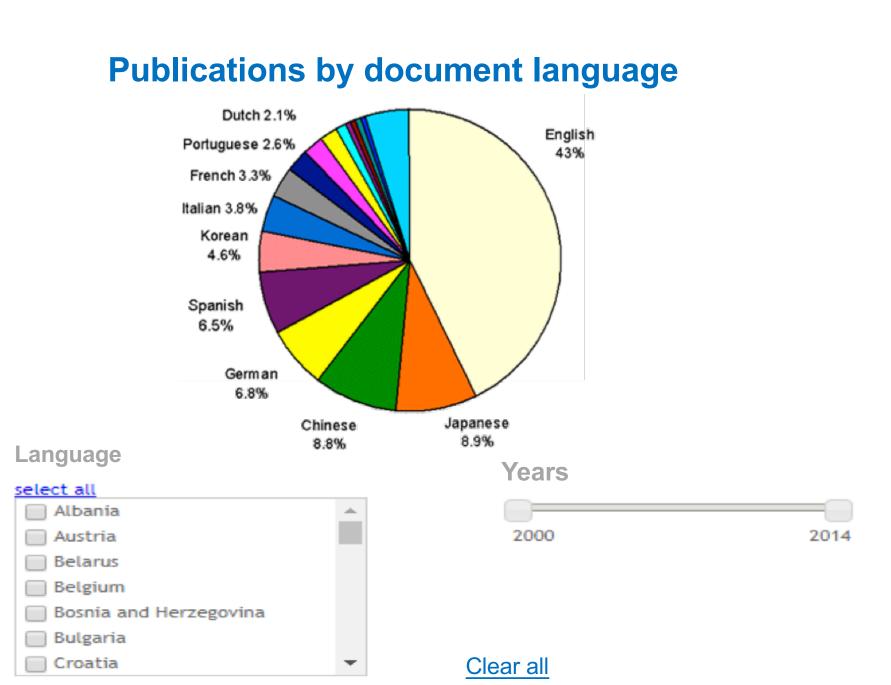
Openness
Publications Distribution
Publications by Scientific area
Publications per Project
Publications by Document Type
Publications by Language



Clear all

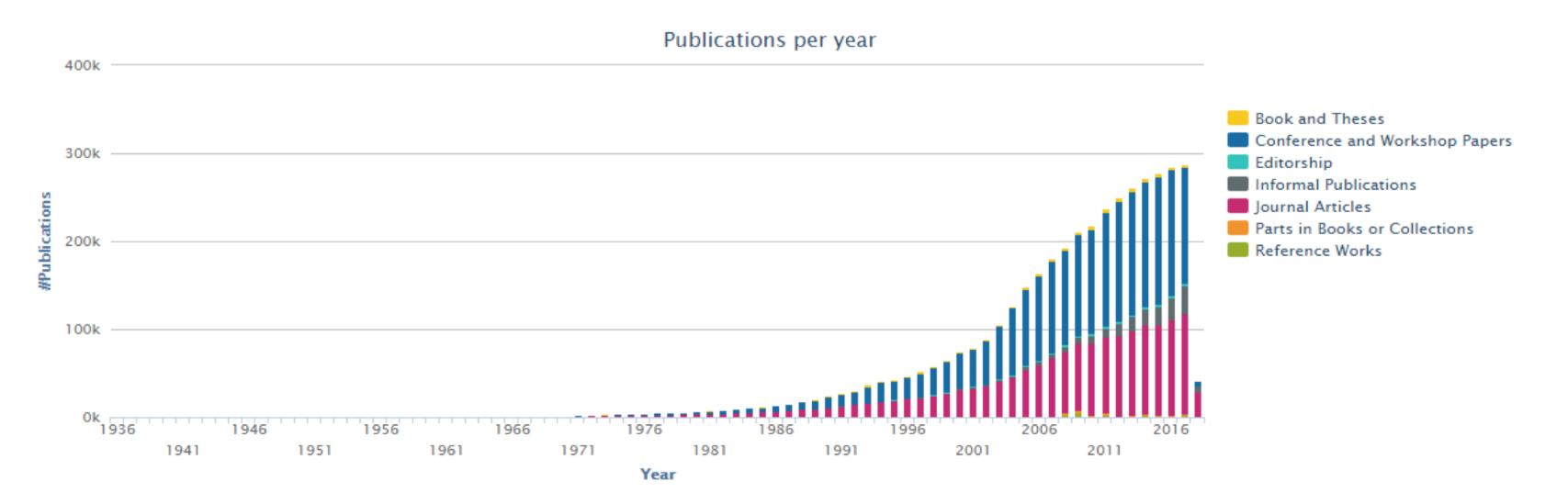


Openness
Publications Distribution
Publications by Scientific area
Publications per Project
Publications by Document Type
Publications by Language



# Combined statistics for publications

# Publications per year by document type







# New products per year by document type

#### Stack bar combined content







# Table combined content

Publications by project per year					
List of projects					
	2017	2016	2015		
Project 1	12	50	47		
Project 2	35	85	68		
Project 3	20	9	23		





# Table combined content

Publications by scientific area by articles, chapter in book and by no. of projects					
Scientific area	Articles	Chapter in Book	No. of Projects		
Bioscience	5	35	2863		
Mathematics	45	20	1852		
Chemical Engineering	2	88	7514		
Medicine	59	14	6528		





# **Table combined content**

Publications by scientific area by articles, chapter in book and by no. of
projects

Scientific area	Articles	Chapter in Book	No. of Projects
Bioscience	5	35	2863
Mathematics	45	20	1852
Chemical Engineering	2	88	7514
Medicine	59	14	6528



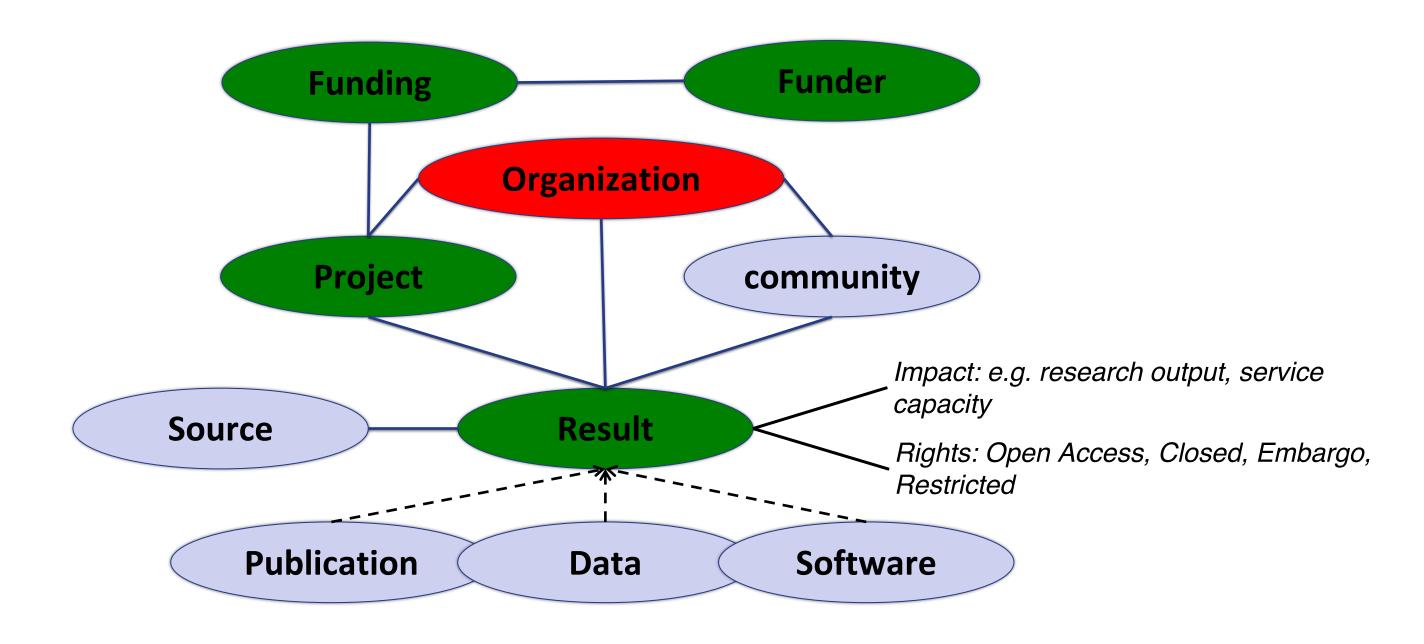


# Institutional Dashboard





# Institutions and OpenAIRE Data Model







## **Aims**

#### Research Impact

Ability of researchers affiliated with the institution to produce innovative and quality scientific products

# Service capacity impact

Ability of services maintained and operated by the institution to support researchers at producing or storing scientific products

# Funding impact

Ability of institution to reach funding from different funders and disciplines





## Added-value functionalities

#### Funders

- Recent and past EC and other funders' activities (representing various funding levels)
- Checking compliance to funder mandates

#### Institutions

Collaboration network (by institution) via projects and products

## Projects

- Check if projects are eligible for Post-Grant APC funding
- Compare project portfolio against that of other similar institutions (anonymized)





## Users

#### Generic users

- Search & Browse
- Visualize Stats
- Claim (registered users)

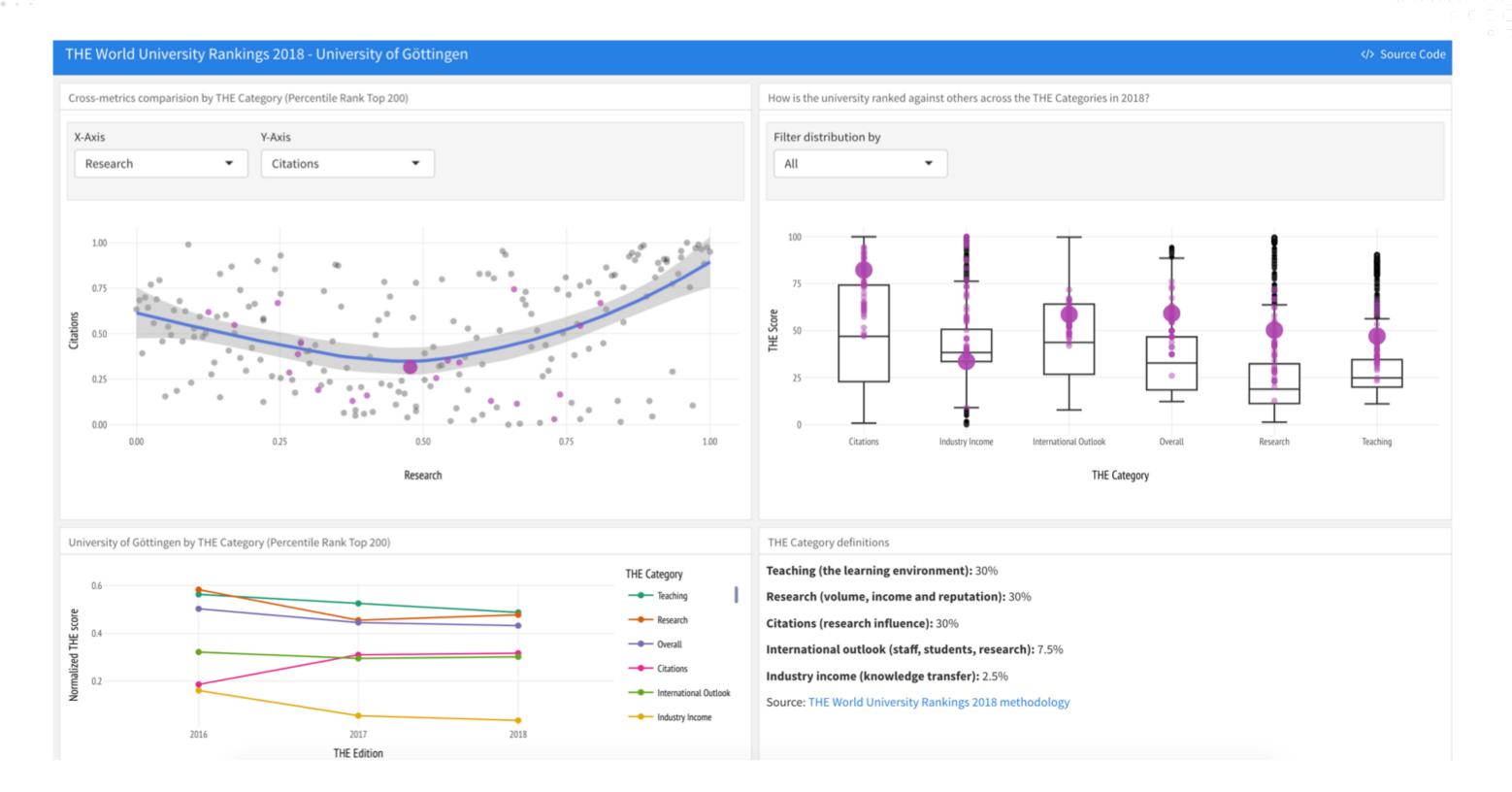
# Managers (registered/authorized users)

- Services to be monitored (harvesting)
- Continuous monitoring (service APIs)
- Alerts (e.g. violation of OA mandates)
- Impact propagation





# Examples of other dashboards







Year





493 - Publications to date, where: ②

26.8% - Verified Publicly Accessible on Publishers Site 2

CHCRUS Dashboard

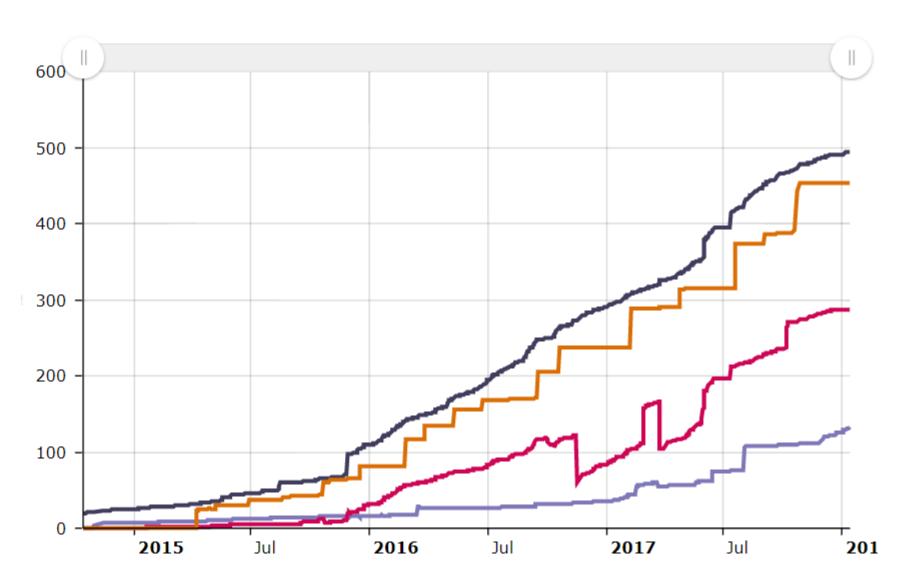
**58.0%** - Reuse terms available ②

**91.9%** - Archived **②** 

#### Key Performance Indicators - History

Click on colored dots in legend to show or hide key indicator lines





#### cran.rstudio.com

The streaming data is a 1-weekdelayed livestream of download logs from cran.rstudio.com. The server code for that service is at jcheng5/cransim.

Warn when rate exceeds:



Recent downloads window:

50

22.9

Downloads per sec (last 5 min)



211

**Total downloads** 



**65** 

**Unique users** 



Popularity by package (last 5 min)

lmtest lme4 leaps ticeExtra\_xpDes mime magrittr sonlite gridExtra pkgconfig knitr e1071 iteratorficTreeCut markdown highr igraph 'nloader Parallel dplyr httpuyoMC rgdal stringi stringr evaluate digest rgl .idyselect data.tablgtoolsctlabels gtable BH bindrcpp yaml rlang asserttha qgplot2 curl
R.methodsS3bindforeach geosphere oxme
signal caToowplot sp ibigeepasse6car

Percent of downloads (last 5 min)

	Package	Percent
1	evaluate	2.8
2	digest	2.4
3	highr	2.4
4	markdown	2.4
5	stringr	2.4
6	yaml	2.4
7	ВН	1.9
8	bindrcpp	1.4
9	data.table	1.4
10	dplyr	1.4
11	glue	1.4



Based on UGOE EC-funded research activities, we will

- Collect use-cases and analytical demands for institutional dashboards
- Explore how and to what extent OpenAIRE data streams (REST API) can be meaningfully represented for analytical purposes at academic institutions
- Create a mock-up with interactive visualisations and reporting functionalities
- Plan a stakeholder testbed to prove dashboard's usefulness and effectiveness with research administrators, and other stakeholders including SHARE.

 Use the OpenAIRE Network for the testing phase, and the NOADs to test and promote the dashboard

# How will this happen?





# Questions?