Implementing Open Science in EOSC

Putting the puzzle together

Natalia Manola
OpenAIRE Managing Director
Athena Research & Innovation Center

Paolo Manghi
OpenAIRE Technical Director
CNR-ISTI

Open Science

Open Access to publications
Open / FAIR data
Open Software
Linked Open Science (Provenance)
Open methodology (Open peer review)
Access to resources for analytics
Access by non-academics

… practice science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.

OpenAIRE

open and **reproducible** science

scientific/scholarly **communication**

A key pillar of EOSC

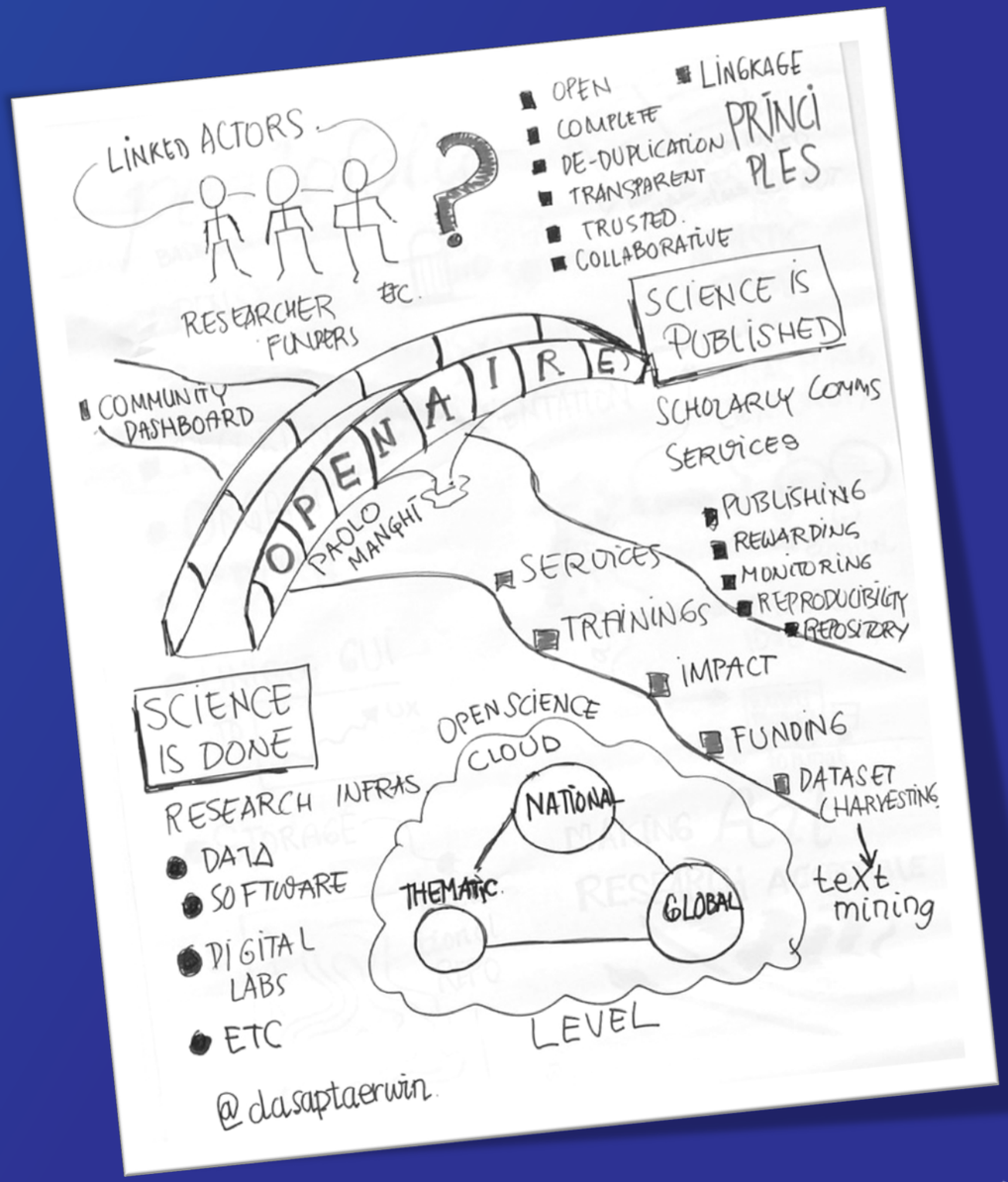
a **infrastructure**

nal + **technical** links

service + data **interoperability**

OpenAIRE

Bridging the worlds where science is performed and science is published



EOSC as a facilitator of Open Science

Actors

?

?

Services

Research infrastructure services, i.e. digital labs (performing science)

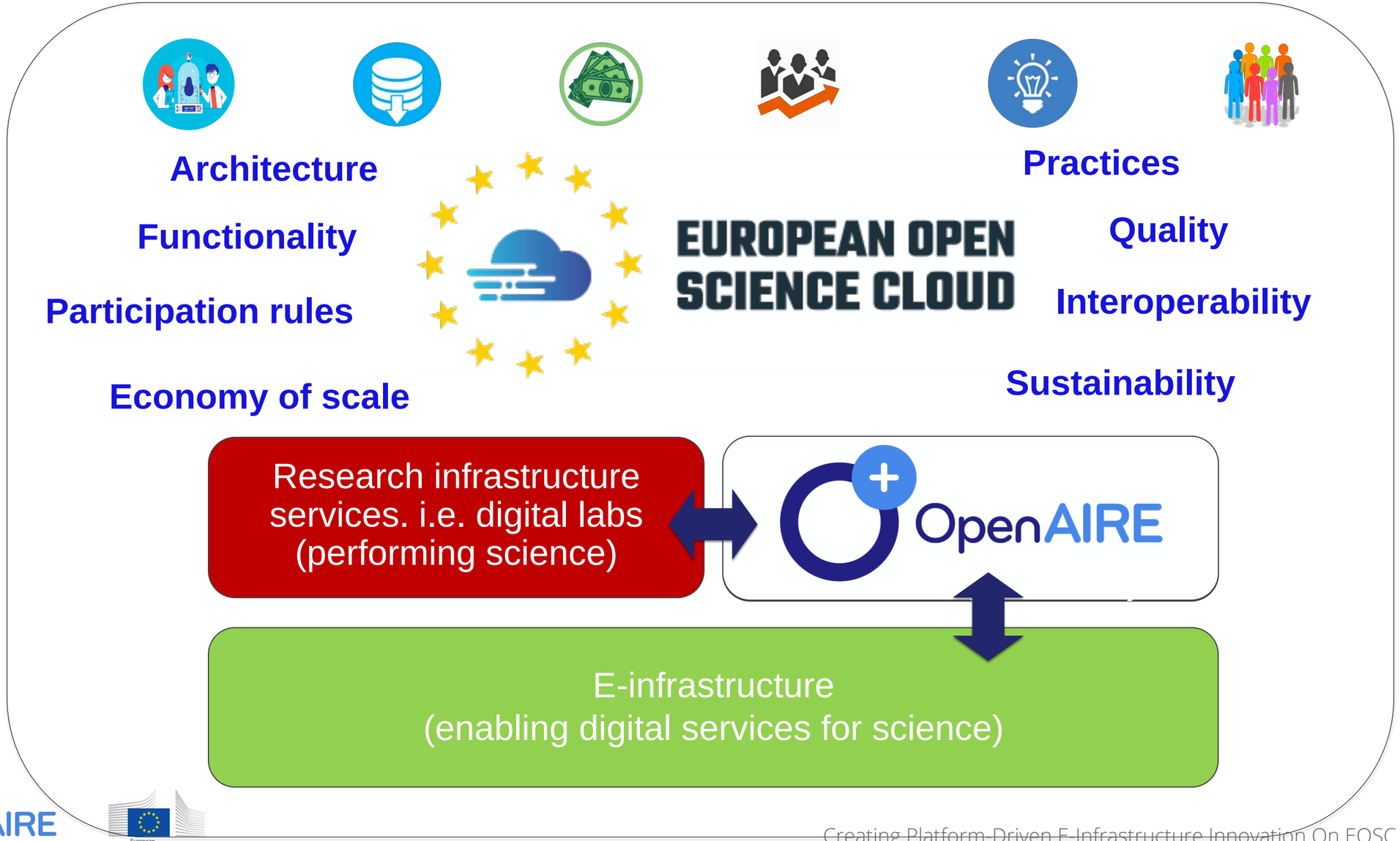
Scholarly communication services (sharing, evaluating, monitoring science)

?

?

E-infrastructure
(enabling digital services for science)

EOSC as a facilitator of Open Science



EOSC, Open Science and data

Small data, Big data

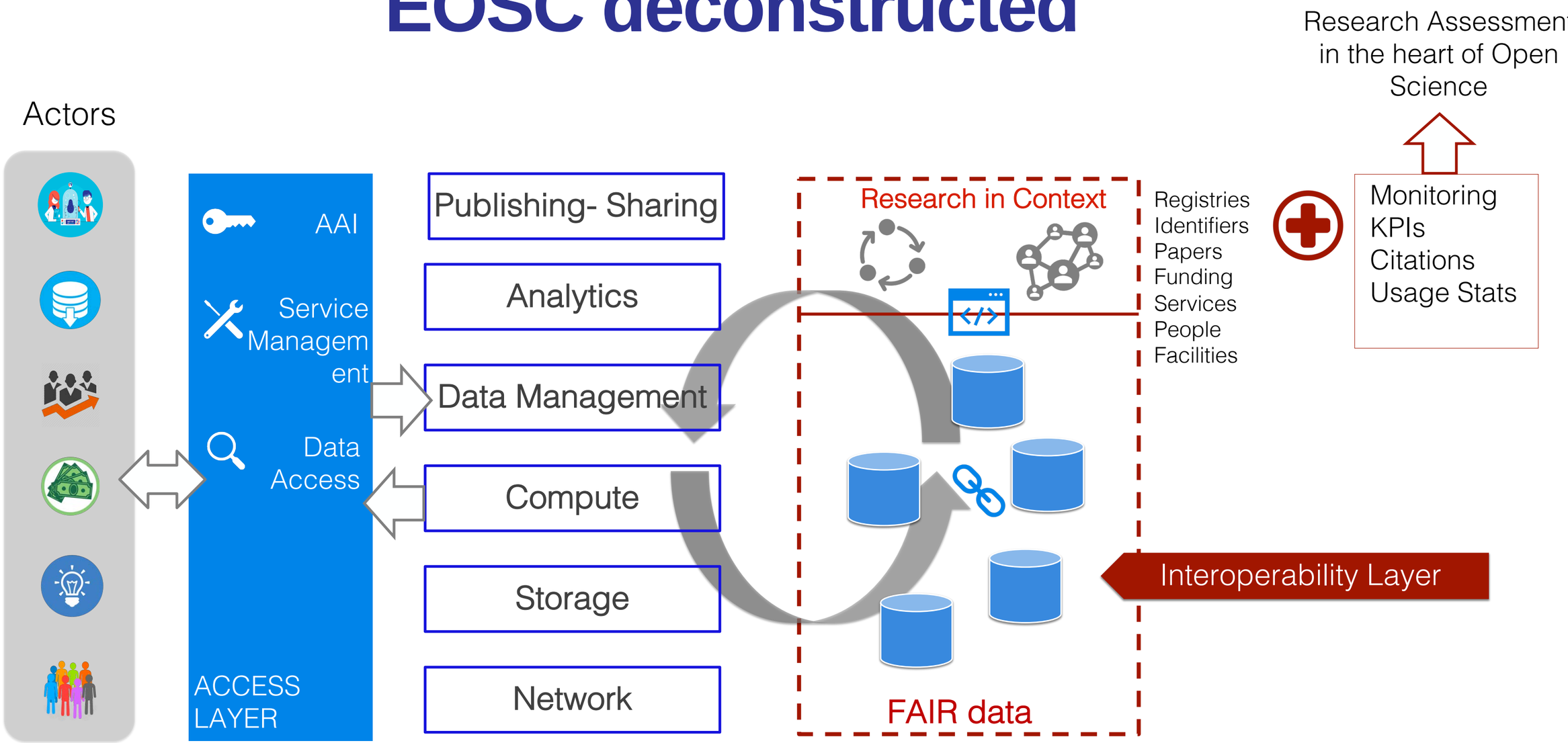
	Small Data	Big Data
Data Source	Accessible, informative, actionable	No traditional data processing
Volume	< 1 TB	Terra and Exascale
Velocity	Controlled and steady data flow	Very fast Speed Fast accumulation
Variety	Structured data	High Variety Data Sets
Veracity	Less noise as controlled collection	Rigorous data validation required before processing
Value	Business intelligence, analysis, reporting	Data Mining for prediction, pattern finding, etc.
Time Variance	Historical data equal valued	In some cases data gets old
Data Location	Databases, local servers	Distributed storages on Cloud
Infrastructure	Predictable resource allocation	Agile Infra, with horizontally scalable architecture

Differences in: Collection, Processing, Scalability, Modeling, Storage & Computation Coupling, Data Science, Data Security

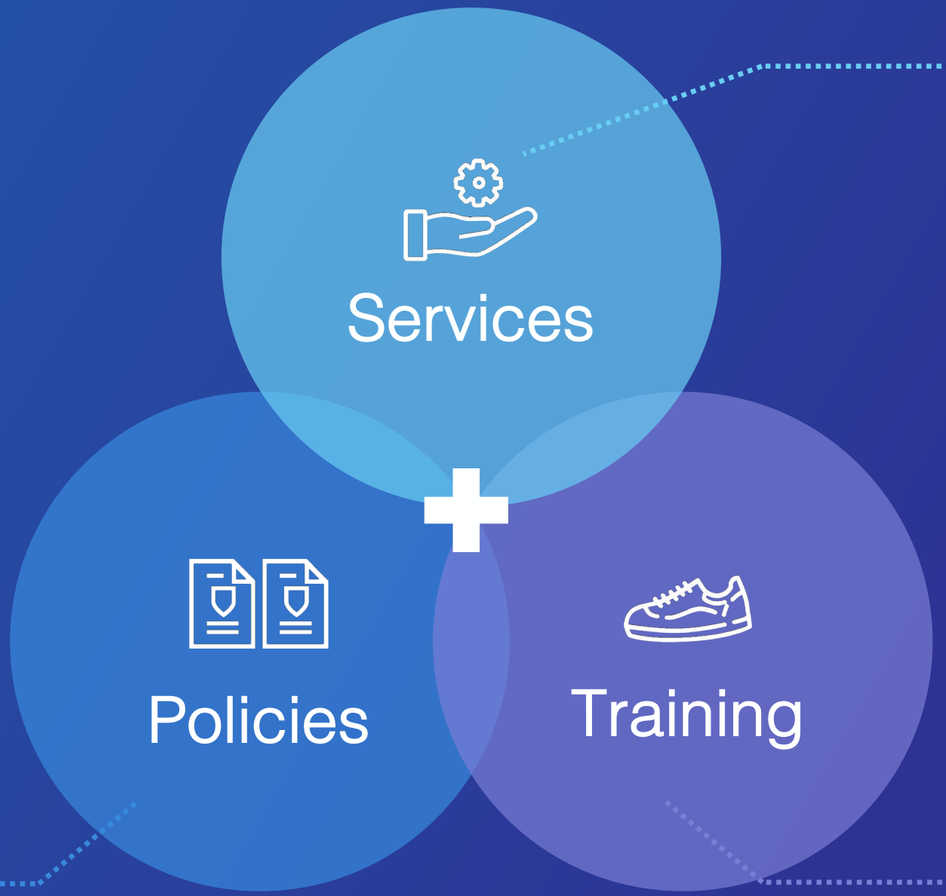
Small data combined needs big data infrastructure

…small data will increasingly be made more big data-like through the development of new data infrastructures that pool, scale and link small data in order to create larger datasets, encourage sharing and reuse, and open them up to combination with big data and analysis using big data analytics

EOSC deconstructed



OpenAIRE 3 pillars of action



Connecting

- *Establishing infrastructure*
- *Added value services*

Aligning

- *Standard*
- *Guidelines*
- *Practices*
- *Workflows*

Empowering

- *Open Science*
- *Open Access*
- *Policies*
- *Services*

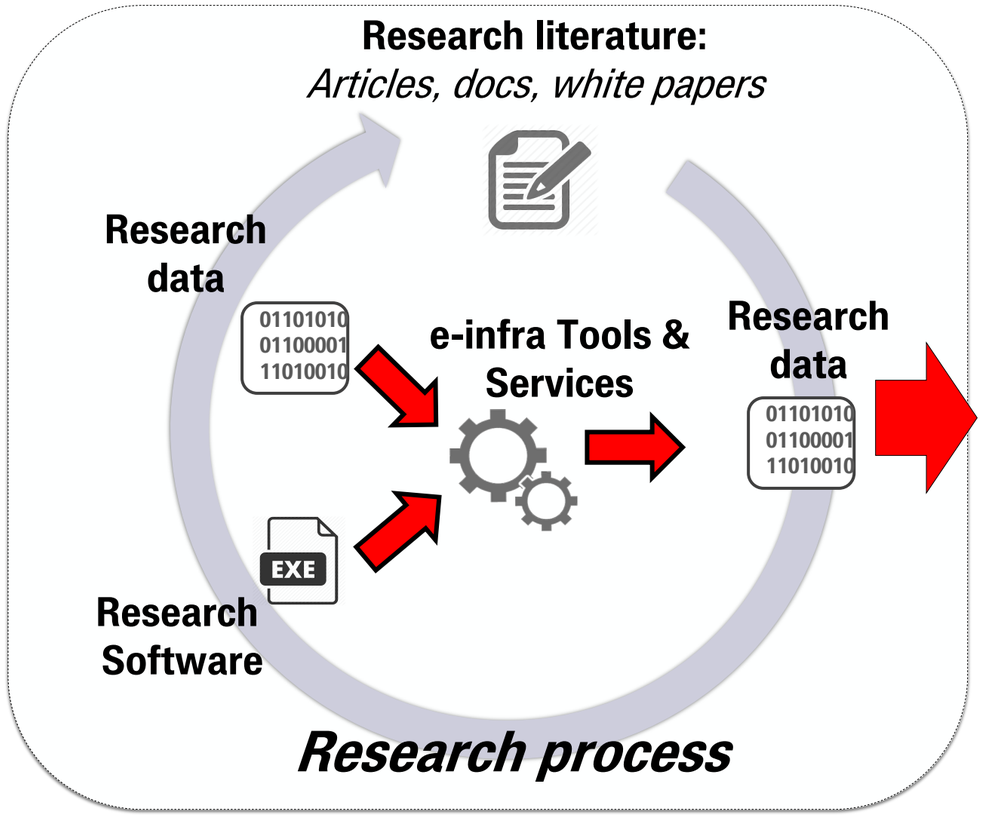
1. Services

Providing the glue via scholarly/scientific communication

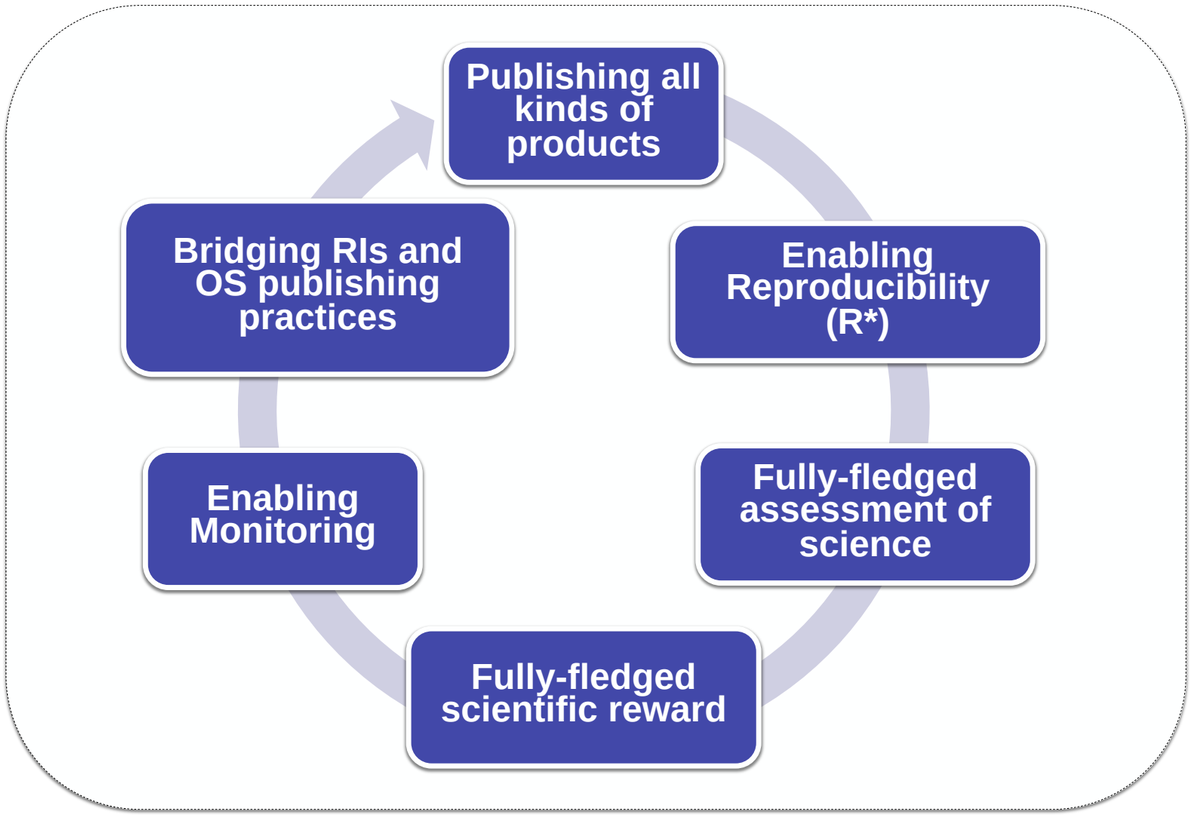
Making small data big

Scholarly Communication transition to Open Science

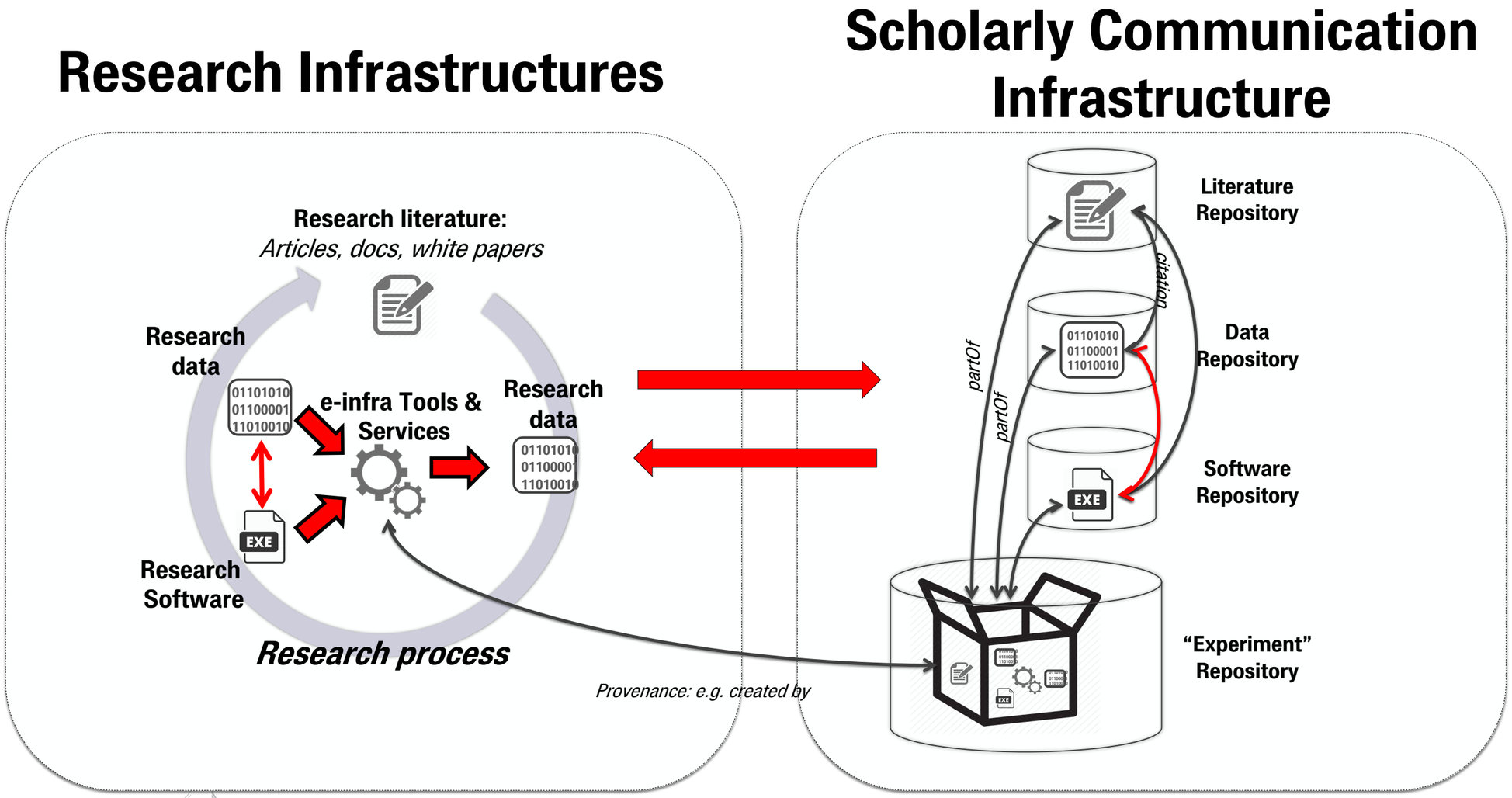
Research Infrastructures



Scholarly Communication Infrastructure



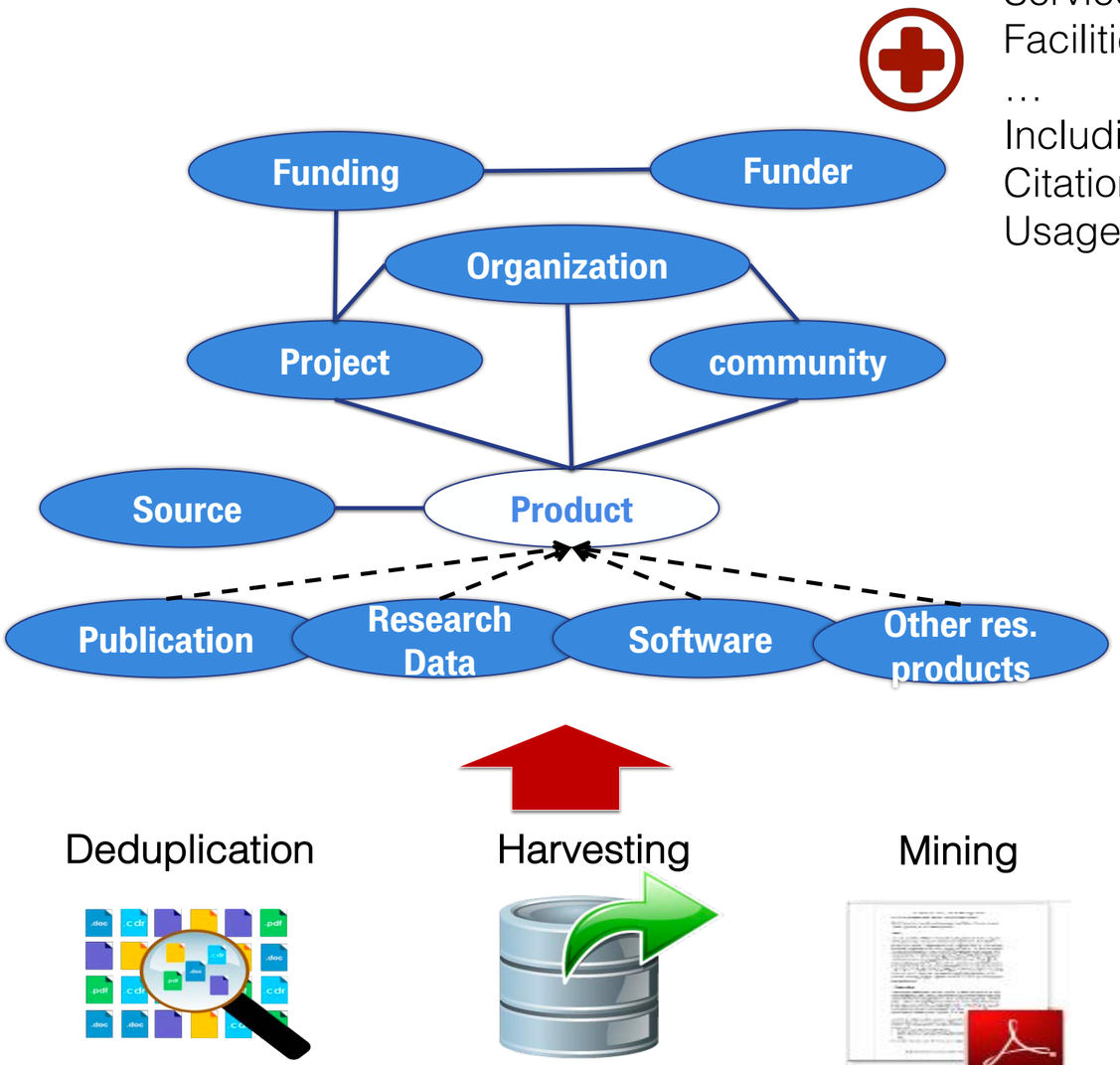
Open Science and Scholarly Communication



OpenAIRE Research Graph

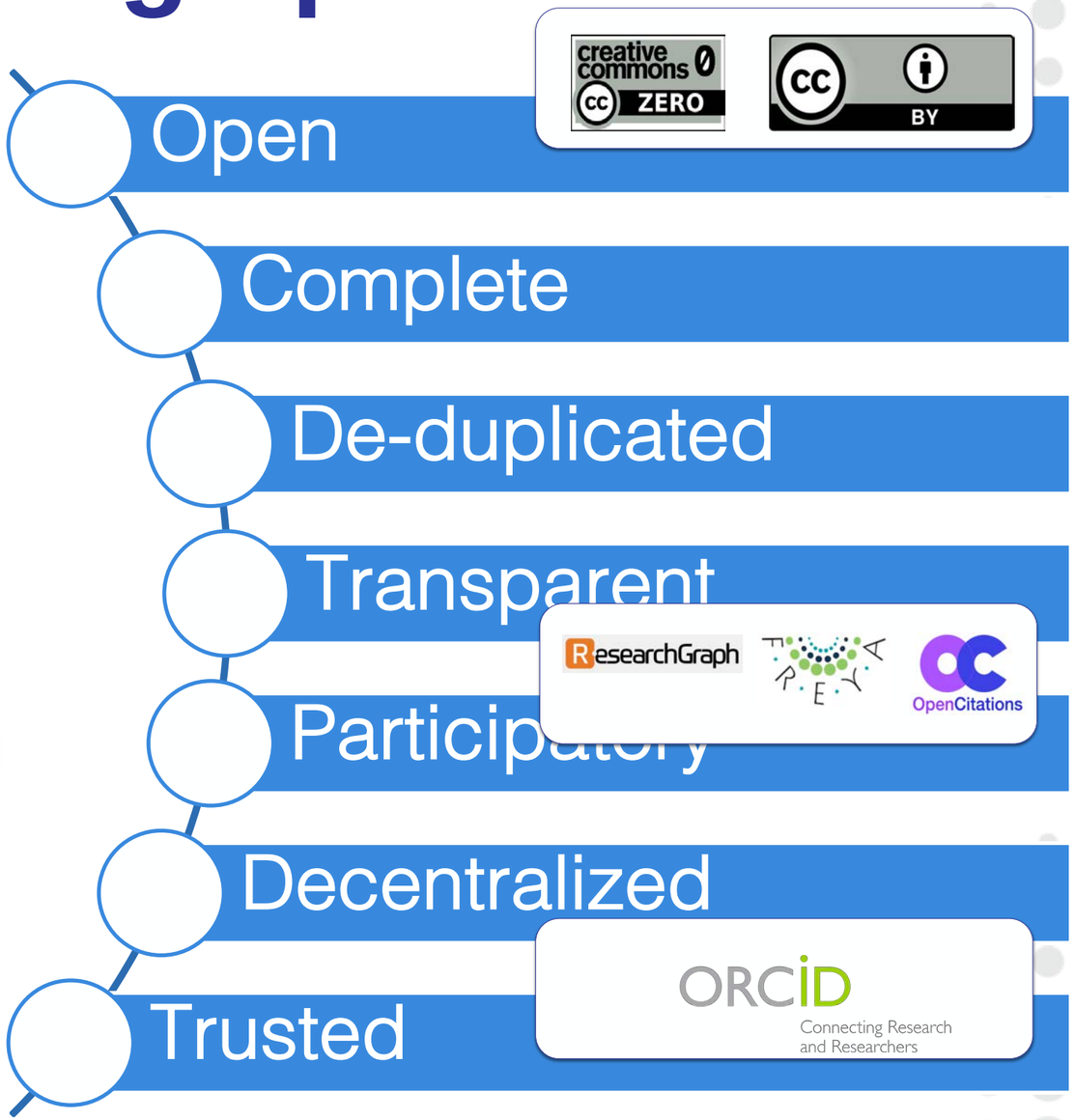
Materializing the Open Research Graph

- Harvested data sources
  - 10K +
- Harvested records
  - 450Mi +
- Publication full-texts
  - 10.5Mi+
- Harvested/mined links
  - 340Mi +



The OpenAIRE research graph

Providing an **open metadata** research graph of **interlinked scientific products**, with Open Access information, **linked** to funding information and research communities



Strategic for Open Science

EUROPEAN OPEN SCIENCE CLOUD

Making the research graph an EOSC resource

Open, Trusted, Complete, De-duplicated, Participatory, Transparent, Decentralized

Actors

Institutions, research organizations, funders, content providers, researchers, SMEs, etc.

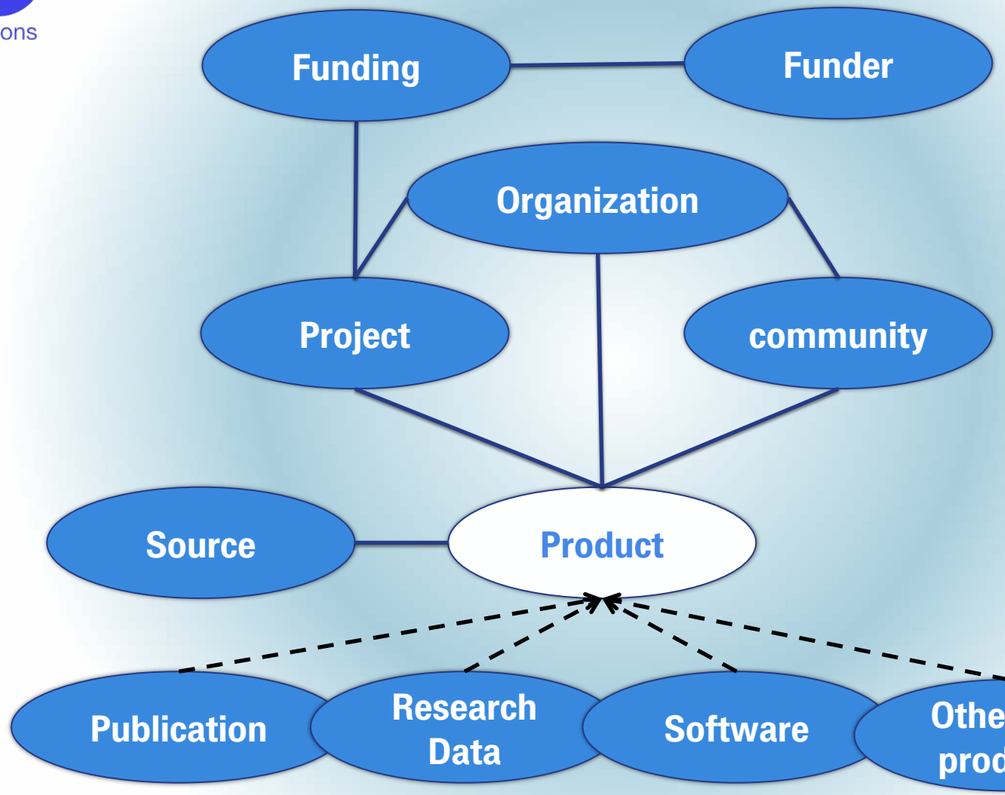
Added value services

Discovery, monitoring, assessment of research

Links to non-academic infras

OpenAIRE

Complete aggregation coverage



… and more

… and more

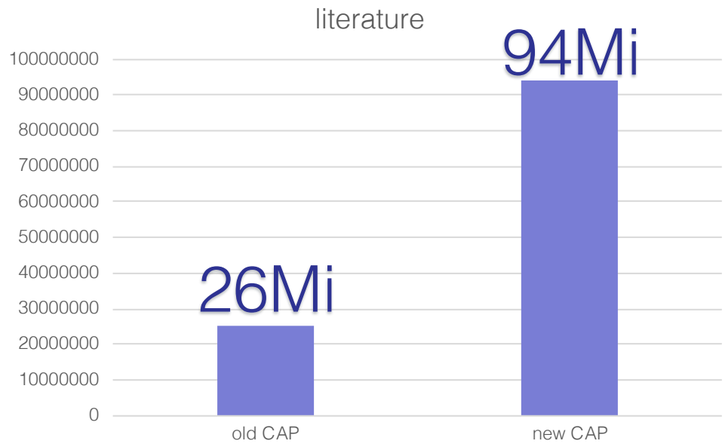
… and more

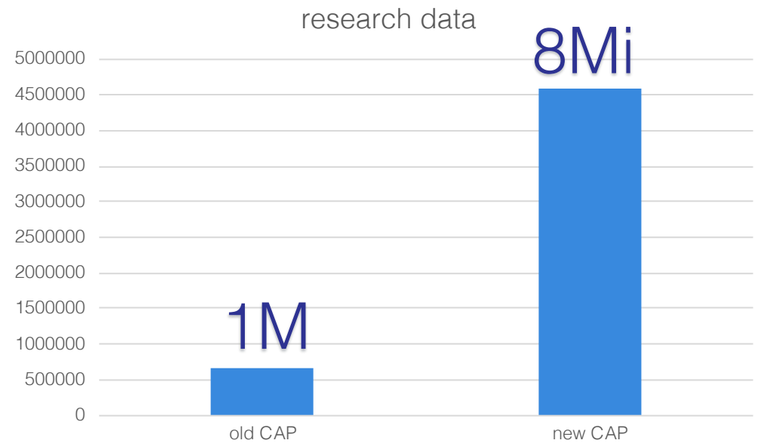
… and more

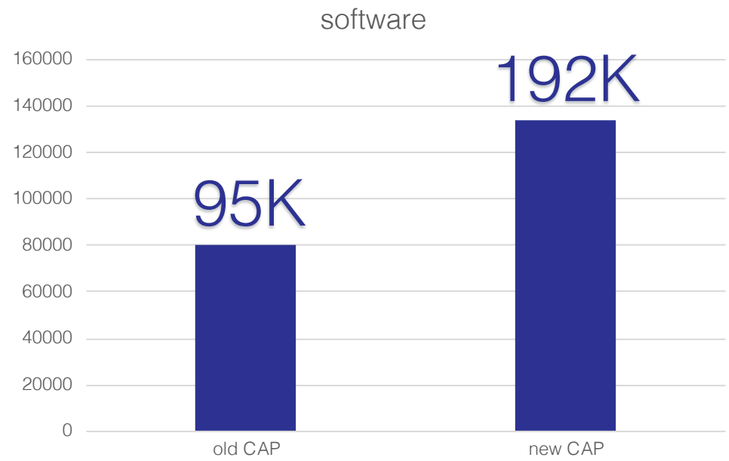
… and more

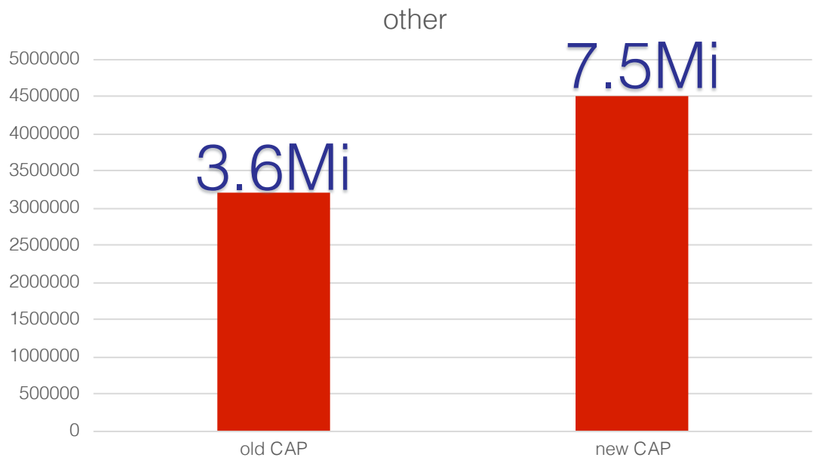
Transition from OA content acquisition policies to OS content acquisition policies

numbers from: explore.openaire.eu and beta.explore.openaire.eu









120Mi

literature-research data links

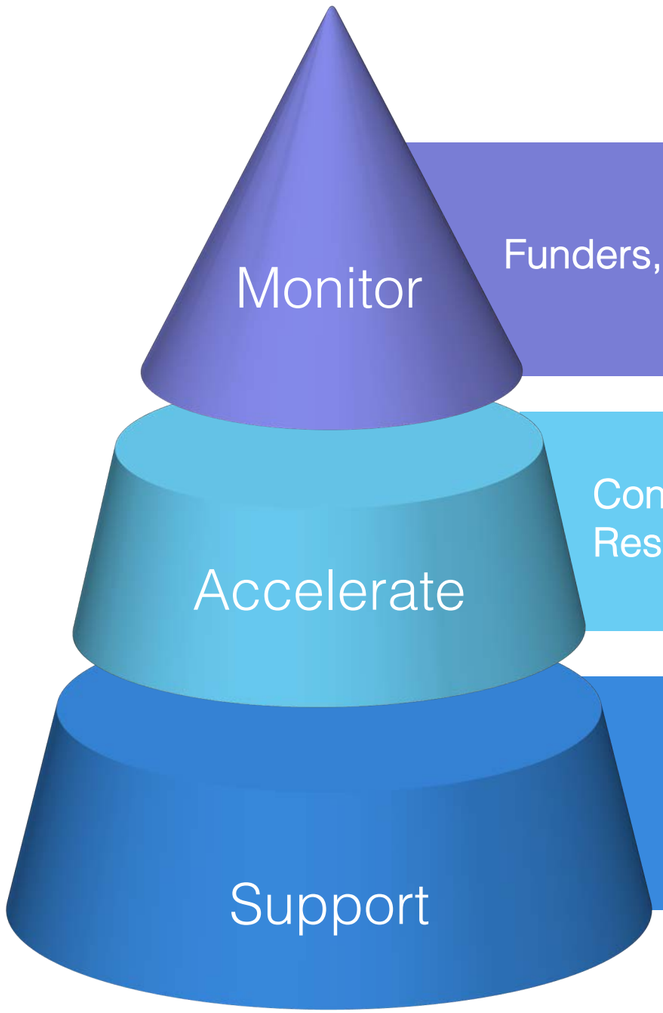
10Mi+

Open Access PDFs for mining

225Mi inferred links:

Article-project
Article-article
Article-software
Article-community
Ecc.

Services for all stakeholders



Funders, institutions, RIs, initiatives, 3rd parties

Research Analytics

Funder Monitoring Dashboard

APIs

Content providers, Research Infras

ScholeXplorer

Interoperability Guidelines

Usage Analytics

OA Broker

Researchers, scientists

Training

zenodo

amnesia

OpenDMP

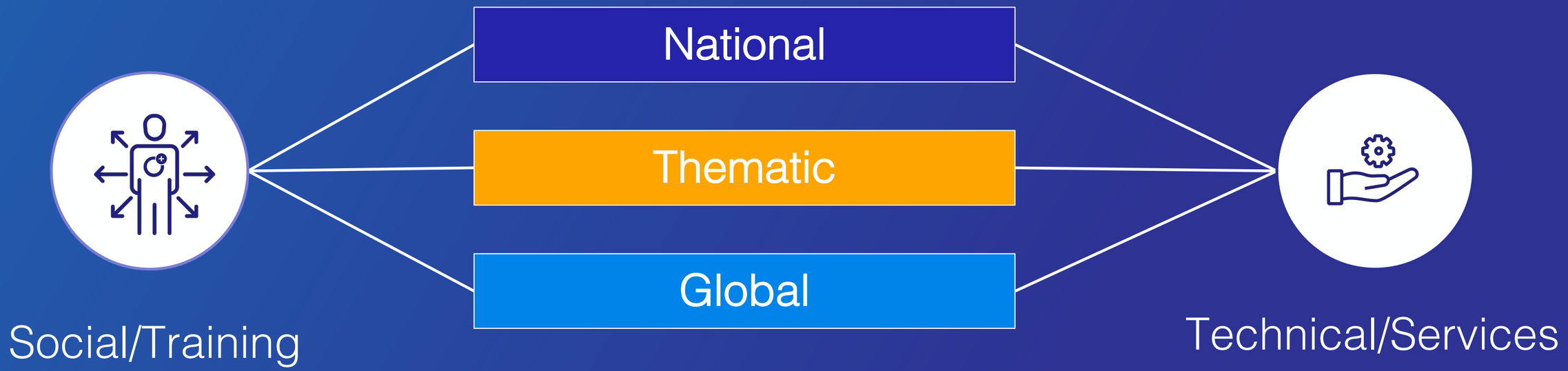
OpenAIRE

2. Support and training

Providing the human aspects

Making the local global

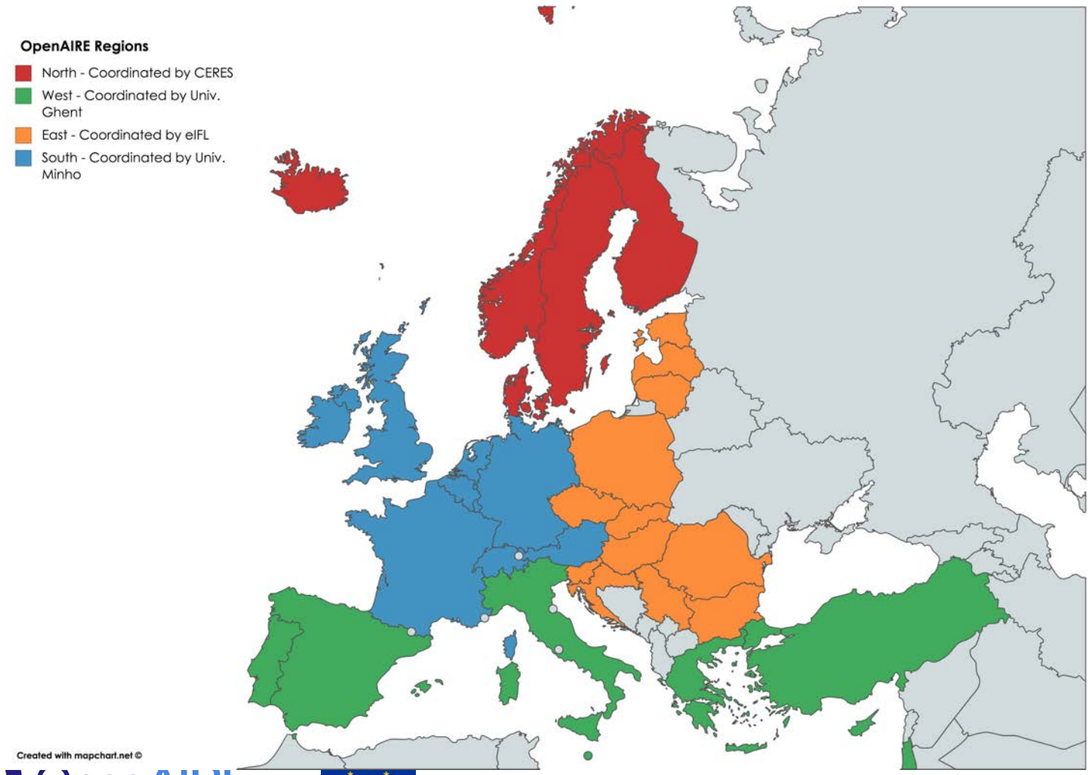
OpenAIRE 3 levels of operation



National Strategy

National Open Access Desks (NOADs)

A pan-European network to address **diversity in culture** & **maturity** of national/local infras



34 countries

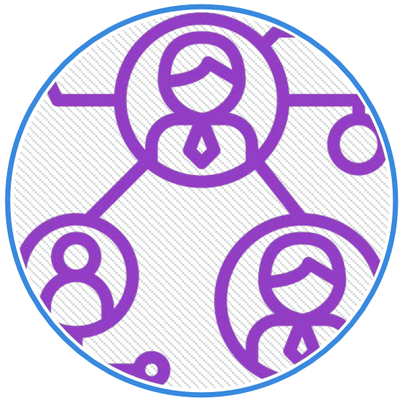
→ Key national organizations

4 regional area coordinators

3 coordinators for

- Policies
- RDM
- Legal

NOADs: A key vehicle in policies and training



Outreach

Support

Training

Policy

10 national workshops **1048** participants
170 conferences attended, presented in **96**
9 funder mandates
4109 repositories, **1720** OA journals contacted

2018

Ground work for OS and EOSC

Support and Training

Distributed and hierarchical training: *train-the-trainers*

NOADs ⇢ National / research infras, organizations ⇢ Researchers

HELPDESK

- Ask a question
- FAQs

RESOURCES

- OA guides
- Copyright issues
- Factsheets

TRAINING

- Webinars
- Workshops

45 webinars **2790** participants

55 f2f training events **1637** participants

8 train-the-trainer events **155** OS trainers

2018

Cross infrastructure OS training

It's all about synergies

- **Rules:** Open Science **policies**
- **Practices:** **Openness** and FAIRness RDM
- **Technical:** **APIs** (ResourseSync, schema.org), **OpenAIRE Guidelines** for Content Providers (metadata)

Community of practice for training the trainers

Thank you!

Natalia Manola

natalia@di.uoa.gr