



EOSC-Hub and OpenAIRE-Advance collaboration: RDA Working Group Incubator

Paolo Manghi

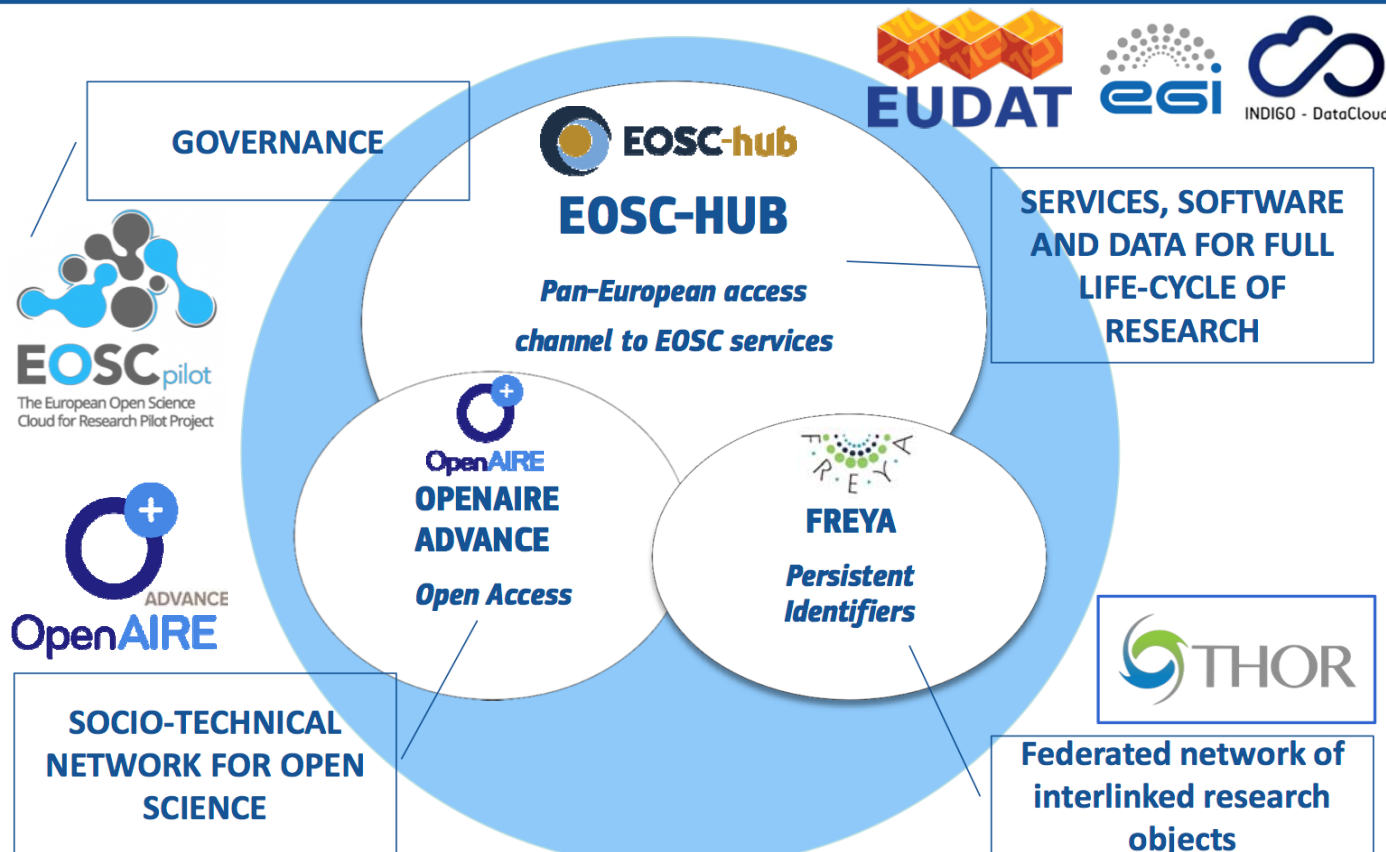
Institute of Science and Information Technologies, CNR
OpenAIRE Technical Coordinator

Acknowledgments to Mark van de Sanden,
SURFsara, EUDAT CDI Technical Coordinator



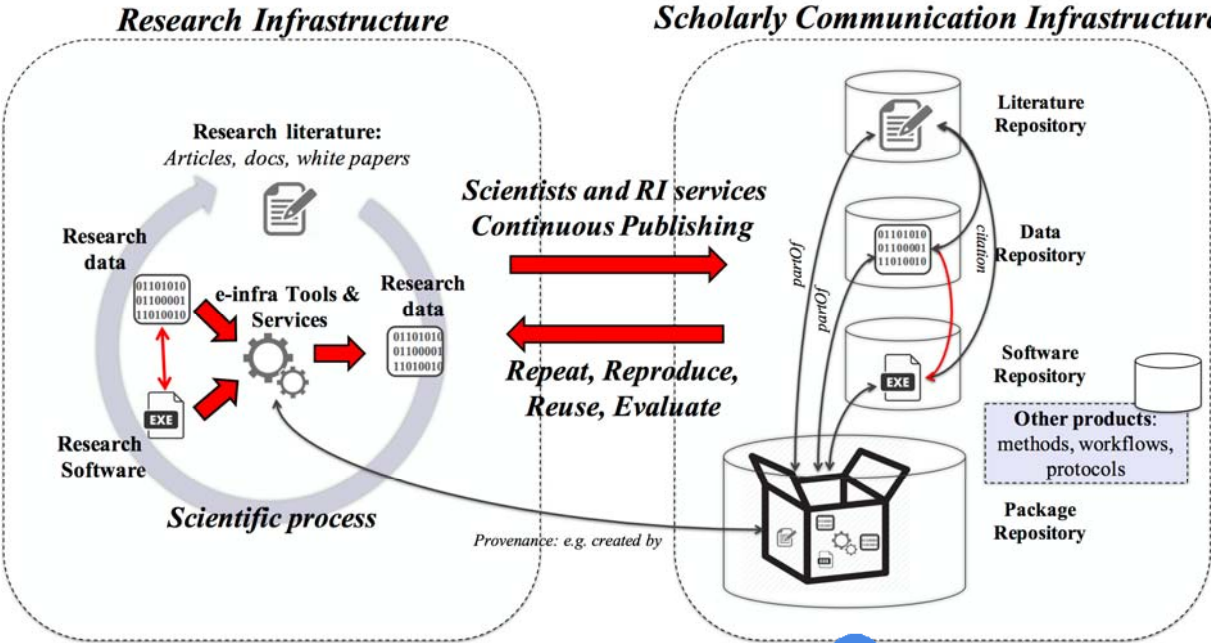
e-infrastructures in EOSC

H2020 Phase: Building and consolidation



Slides: Thomas Skordas, Director, DG Connect, European Commission – presented at DI4R 2017, Brussels

Research and Scholarly Communication



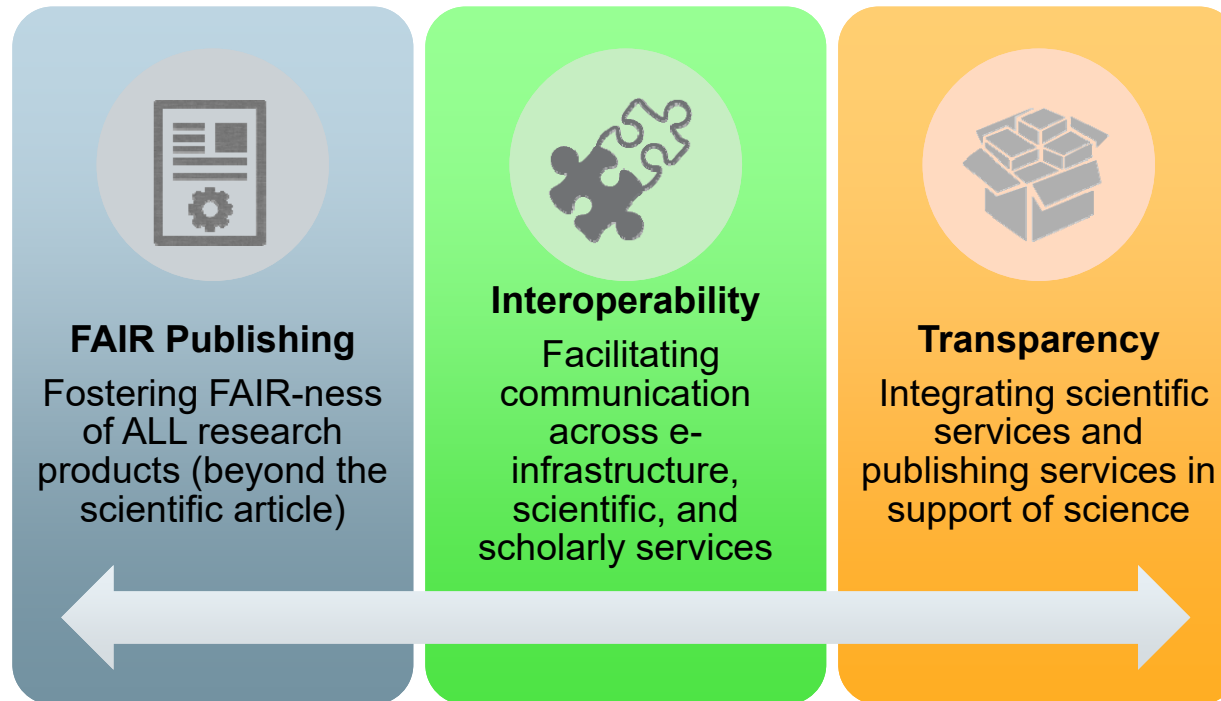
**Storage and Computation:
services and policies**



**Open Science Publishing:
services and policies**

EOSC-Hub and OpenAIRE

Developing scientific and scholarly Open Science commons



Developing synergies and roadmap



Team A: Governance and sustainability

Facilitate access to integrated services at European and National level



Team B: Outreach, Support and Training

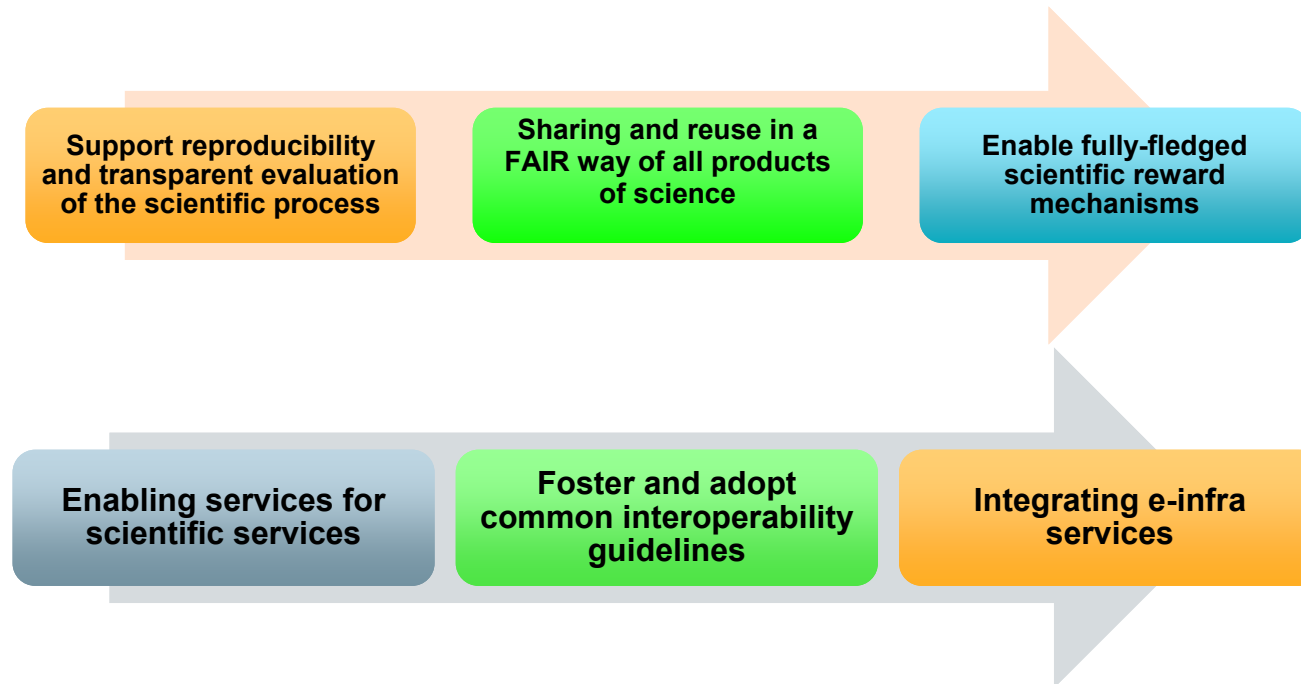
Deliver scientists a common message



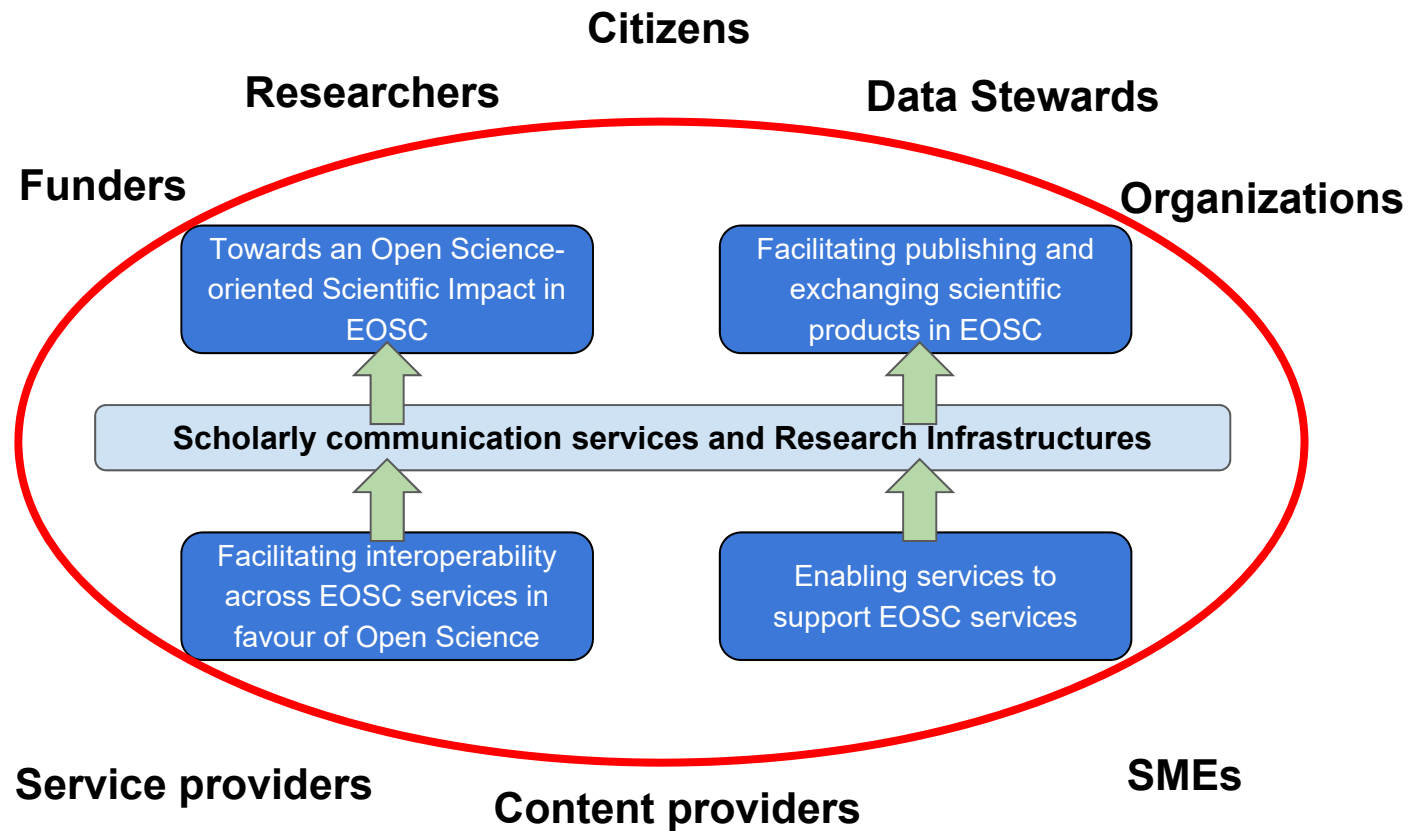
Team C: Service integration

Commons: DMPs, usage statics, guidelines, enabling services

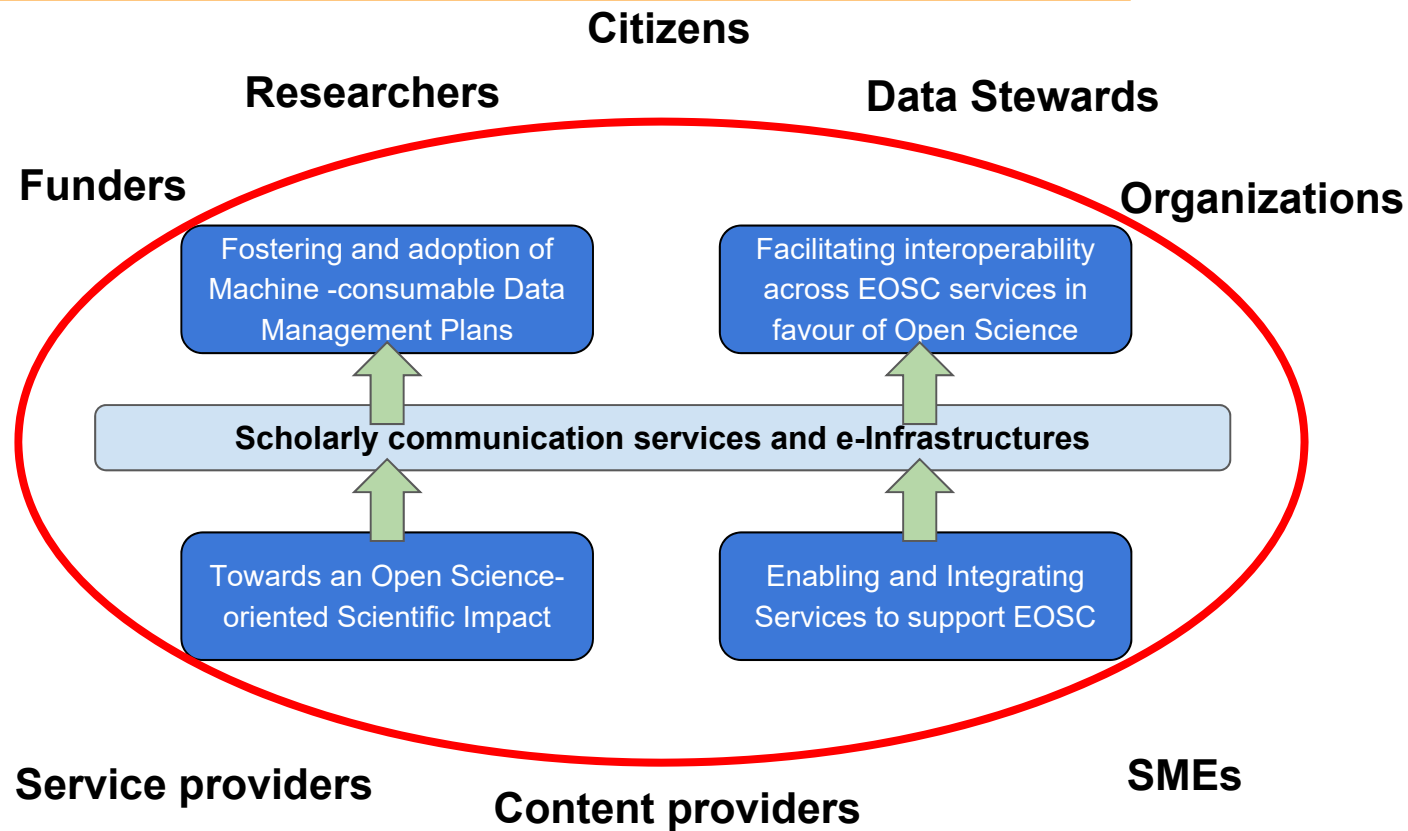
Team C: service integration



EOSC Service Integration



Team C: service integration



Fostering and adoption of machine-consumable Data Management Plans

Collaboration
with DMP
RDA WG

Pilot DMP
tool with
communities

Explore
publishing of
DMPs

DMP tool as
a catalogue
product

Fostering and adoption of Machine - consumable Data Management Plans

Type of output

Service

Stakeholders: researchers, data stewards, content providers, funders to get advanced DMP management and relative statistics



Benefits

Researchers and data stewards have better and easier methods and tools to describe data management plans in relation to needed requested services and resources to support planned scientific research

Service providers can automate the demand and supply chain between requestor (researcher/community), funder and service provider and optimise capacity management

Funder and research communities have easier ways to assess, reward and and verify scientific grants throughout the whole lifecycle

Facilitating interoperability across EOSC services in favour of Open Science

Define and promote common guidelines for scientific product content providers

Facilitating publishing and exchanging information on scientific products in EOS

Foster guidelines for data archives

Draft guidelines for software and other products

Define guidelines for communities

Adapt EOSC-Hub services to guidelines for other products

Make EOSC-Hub services interoperate with OpenAIRE

Pilot OpenAIRE brokering services with EOSC-Hub Services

Define and promote common guidelines for scientific product content providers

Type of output

Specifications for metadata exchange

Stakeholders: content providers (data providers, software providers, publishers) to maximize accessibility, interoperability, findability; SMEs which can count on standards for metadata exchange

Specifications for usage stats sharing

Stakeholders: content providers, researchers, organizations, funders for research impact evaluation of their products (definition of new citation/quality indexes for science)

Specifications for community identification and relationships

Stakeholders: researchers, service providers to provide and consume community-flavoured services

Benefits

Researchers and data stewards have clear guidelines to describe research/scientific products (research data, software, experiments, objects, etc.)

Research communities have clear guidelines to make/adapt thematic services FAIR (Findable, Accessible, Interoperable and Reusable) to increase scientific impact of research output

Content and service providers have clear guidelines to make/adapt repositories for scientific products (publications, data, software and other) FAIR to increase scientific impact of research output

Facilitating publishing and exchanging information on scientific products in EOS

Adapting EOSC services to guidelines

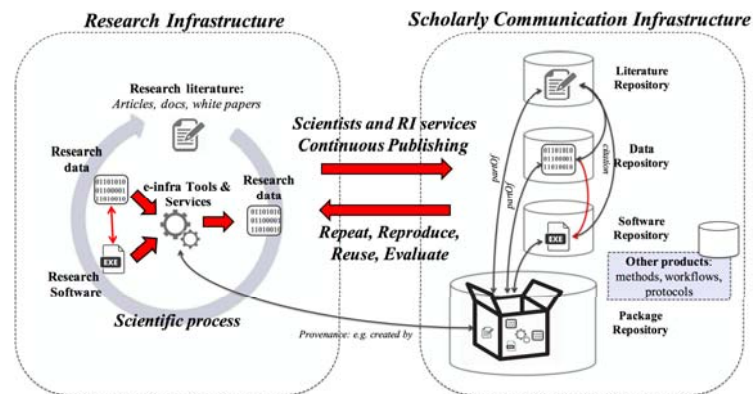
Stakeholders: researchers while performing their scientific process using EOSC-Hub services can (i) implicitly/automatically publish and report their scientific products to the funders while (ii) sharing their products within their community

Extend EOSC services with annotation

Stakeholders: researchers share annotations with each other; service providers and SMEs can benefit from such content to provide useful services

Sharing of scholarly communication

Stakeholders: content providers which will complete, keep up-to-date, enrich their collections in (almost) real-time and the researchers accessing such content providers



Towards an Open Science-oriented Scientific Impact

Define
guidelines
for sharing
usage
statistics in
research
data

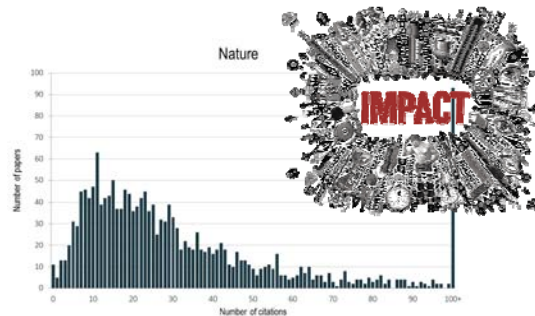
Pilot usage
stats
guidelines
with EOSC-
Hub
Services
and
OpenAIRE

Towards an Open Science-oriented Scientific Impact

Type of output

Service

Stakeholders: researchers can access scientific products together with their usage stats, scientists authoring scientific products different from literature can benefit from new measures of quality; service providers and SMEs can define quality metrics and relative tools for open science (taking into account all products and their usage stats)



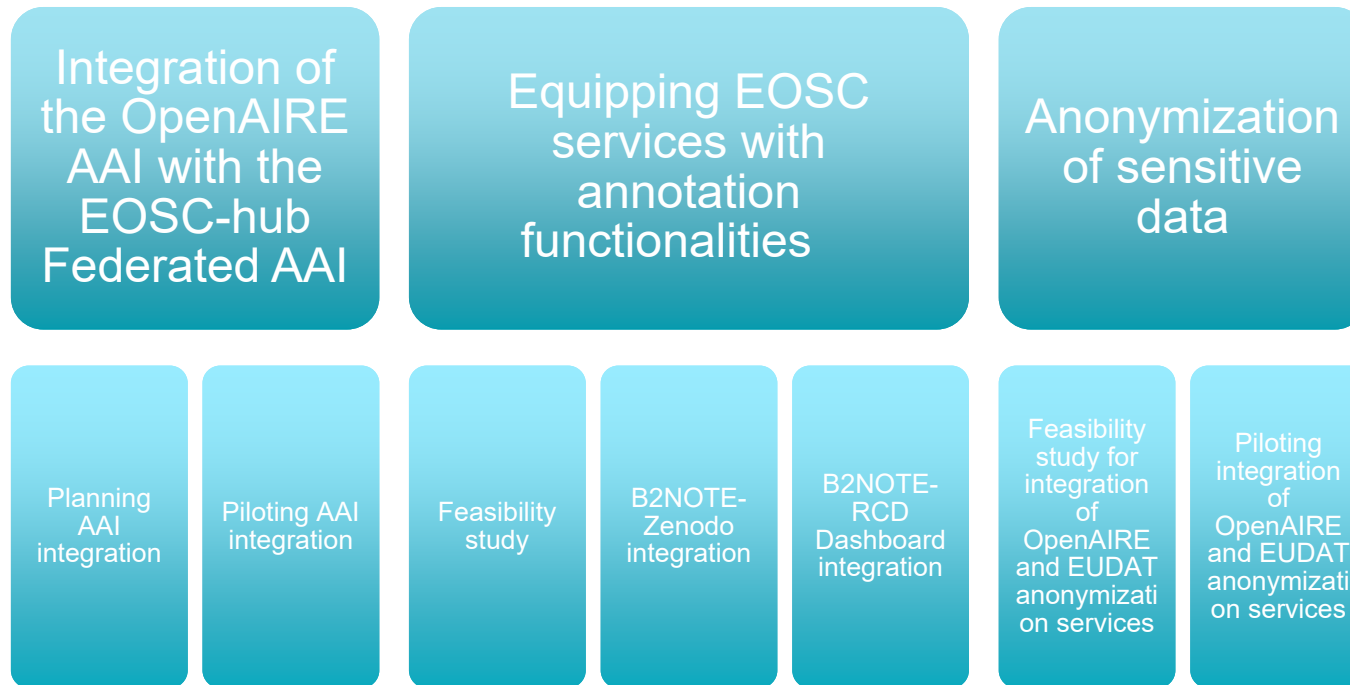
Benefits

Researchers and data stewards have direct feedback and scientific impact on research/scientific output (research data, software, experiments, objects, etc.) produced or can use it as a selection mechanism

Research communities and funders have metrics on scientific impact on research done and funded. It provides metrics within the decision making process for long term preservation of scientific products.

Content and service providers have metrics on the usage and re-usage of content and services, including metrics for long term preservation of scientific products or to sustain services

Enabling Services in support of EOSC



BoF: RDA working group ideas



Community
framework



Usage statistics for
data



Guidelines for “other
products” of science

Community framework

- Define community
 - Define by disciplines, governance?
 - Relationships between communities
 - Relationships with research infrastructures
- Build an authoritative catalogue of communities
 - DOI for communities?
- Use-cases
 - Link RIs, tools, services and products to one or more communities
 - Other catalogues should take the notion of community into account, i.e. service catalogues (e.g. eInfraCentral, EOSC-Hub), product catalogues (e.g. OpenAIRE), etc.
-

Usage statistics for data

- Identify models/standards for collection (interception) of actions/events over data
 - Definition of relevant actions to be tracked: see examples in the publications world, engage/re-use MakeDataCount and similar activities
 - Model: action types, tracking underpinning facilities, etc.
 - Provide/customize scripts for usage events interception (e.g. Lagotto)
- Identify and implement use-cases involving several data repositories and build an aggregation of usage stats
 - See work performed in OpenAIRE, IRUS-UK, and others
 - Deliver sample aggregated statistics

Guidelines for “other products” of science

- **Motivation:** Guidelines for exchanging citation metadata on publications, datasets, and software have been made available Other products, although relevant for science, are left out of this classification: e.g. Virtual Appliances, Workflows, Protocols
- Identification of common citation metadata properties for “other products”
 - See work done on OpenAIRE
 - Single out other “critical mass” classes of products
 - Identify data sources and pilot their compliancy with guidelines



Mark van de Sanden
mark.vandesanden@surfsara.nl

