Paolo Budroni – Martin Semberger

Shaping EOSC - Open Data and the European Data Strategy

Ein Workshop zur Umsetzung der Open Data und PSI Richtlinie der EU (Public Sector Information Directive) und zur Etablierung der European Open Science Cloud in Österreich

TU **Bibliothek›**
WIEN Your knowledge hub

Bundesministerium
Digitalisierung und
Wirtschaftsstandort

Part ONE – Paolo Budroni

- Shaping EOSC

- The European Digital Single Market, Open Science and EOSC

- Regulatory frames

- The added value

Part TWO – Martin Semberger

- The Open Data and Public Sector Information (PSI) Directive

- Implementation in Austria (Informationsweiterverwendungsgesetz)

- The EU Data Governance Act (DGA)

- European Data Spaces

Shaping EOSC

Open Science and the European Open Science Cloud

The European Commission is promoting the European Open Science Cloud. According to the Commission "the EOSC is not an actual cloud service, it is a kind of reengineering of existing e-infrastructures based on scientific data.

The EOSC is a federated environment for the sharing and re-use of scientific data, based on existing and emerging elements in the Member States, with lightweight international guidance and governance and a large degree of freedom regarding practical implementation".

In this context Open Science is seen as a movement to make scientific research, data, and dissemination accessible at all levels of an enquiring society. This vision will be possible only if we understand how to manage and steer the different components and players, at all levels of the forthcoming Open Science Cloud.

- See: http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

Guiding concepts

EOSC and Science

- Key target group: researchers and research support units

- Including research data, educational resources, cultural heritage data

- Bringing science closer to society to address societal challenges

- Fostering trans-disciplinarity

- Widening the scope of services to include <u>users from the public sector</u> and <u>industry</u> as well as <u>citizens</u>

> Inclusion of actors from multiple communities

Guiding concepts

EOSC is not a project – it is a process

Federation of infrastructures and services: Overcoming the current fragmentation by integrating existing infrastructures and building new ones - not a new dedicated infrastructure or software - access for humans and **machines**

Gives a strong push in Europe towards a **culture of data**

It is a **process** of making research data in Europe accessible to all researchers under the same conditions of use and usage
> infrastructures are connected via standards, e.g. for authentication, participation rules, and guiding principles (e.g. FAIR Principles)

A trusted virtual environment for sharing scientific and cultural heritage data, for reusing data, for providing seamless services

The relevance of **legal frames**

EOSC - Some data and computing infrastructures in Austria in Austria

- ARCHE (A Resource Centre for the HumanitiEs)

- GAMS (Humanities' Asset Management System)

- CLARIAH-AT is the consortium of Austrian universities and research institutions that coordinates and drives Austrian activities in the European ESFRI research infrastructures CLARIN (Common Language Resources and Technology Infrastructure) and DARIAH (Digital Research Infrastructure for the Arts and Humanities).

- Phaidra (serving UNIVIE, FWF, Arts Universities)

- AUSSDA (The Austrian Social Science Data Archive) is participating in CESSDA ERIC (Consortium of European Social Science Data Archives)

- BBMRI.at is the Austrian node of the BBMRI-ERIC (pan-European Biobanking and BioMolecular Resources Research Infrastructure with statutory seat and executive management office in Graz)

EOSC - Some data and computing infrastructures in Austria in Austria

- CCCA Data Centre (Climate Change Centre Austria)

- EODC (Earth Observation Data Centre)

- Institutional repositories for publications and/or data at Austrian universities, IIASA, IST, LBI for HTA, Vorarlberger Landesbibliothek, etc.

High performance computing and cloud computing infrastructures

- VSC: The Vienna Scientific Cluster is a collaboration of several Austrian universities that provides supercomputer resources and corresponding services to their users, it is operated by TU Wien.

- MACH-2: The massively parallel shared memory supercomputer MACH-2 is operated in the frame of a project by the JKU Linz on behalf of a consortium of several Austrian universities and research institutions.

- EODC (Earth Observation Data Centre): Cloud computing services

EOSC – Strategic and Innovation Agenda (SRIA)

- Stage 1 (2021–2022): **Development towards added value from a functional federation of infrastructures**

  Enabling the European Open Science Cloud operations (the EOSC-Core) **to provide necessary core functions of the Minimum Viable EOSC (MVE)** that allow federation of existing and future infrastructures, with associated rules of engagement and governance that provision growth and expansion in the following stages.

- Stage 2 (2023–2024): **Expansion to production that generates added value**

  Expanding and building the core data infrastructure **to support the full cycle of workflows for scientific research** in key thematic areas. During this period, work to build on pilots/demonstrators and to link EOSC beyond the research communities to the wider public sector and the private sector will begin.

- Stage 3 (2025–2027 and beyond): **Expansion to develop impact from Open Science**

  Deployment of federated research infrastructures for European researchers with functionality that provisions **actors from multiple communities** to deliver impactful Open Science. In addition to European infrastructures, the national research infrastructures delivered from the Member States and Associated Countries in particular will help in this expansion phase.
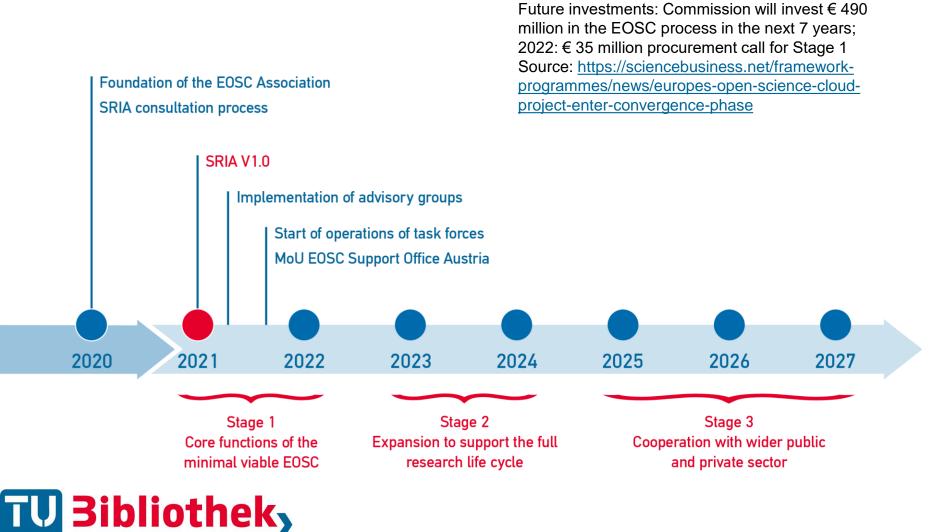
Source: EOSC-SRIA V1.0, p. 150

TU Bibliothek,
WIEN Your knowledge hub

OUTLOOK

Future investments: Commission will invest € 490 million in the EOSC process in the next 7 years; 2022: € 35 million procurement call for Stage 1
Source: https://sciencebusiness.net/framework-programmes/news/europes-open-science-cloud-project-enter-convergence-phase

Foundation of the EOSC Association
SRIA consultation process

SRIA V1.0

Implementation of advisory groups

Start of operations of task forces
MoU EOSC Support Office Austria

2020 2021 2022 2023 2024 2025 2026 2027

Stage 1
Core functions of the minimal viable EOSC

Stage 2
Expansion to support the full research life cycle

Stage 3
Cooperation with wider public and private sector

TU Bibliothek
WIEN Your knowledge hub

SRIA - KPIs

Key Performance Indicators according to version 1.0

Extract

(2023)   70% of **publications** from EOSC Association research-performing members that become **immediate open access**

(2023)   60% research disciplines that have **documented standards and protocols** for data sharing and reuse

(2025)   5 national education systems that recognise **European curricula for data stewardship**

(2025)   50% **research data** from EOSC Association members which is deposited in repositories that is made **as open as possible**

(2025)   70% **metadata** related to publicly funded research datasets which are defined as Open Data that are **discoverable through EOSC** federated **infrastructure**

(2025)   50% of EOSC Association members that **recognise Open Science activities** in research career assessments

(2025)   70% research-funding members of the EOSC Association that require data sharing and incentivise reuse

Source: EOSC-SRIA V1.0, p. 160 f.

Shaping EOSC - The European Digital Single Market, Open Science and the European Open Science Cloud

TU WIEN **Bibliothek**
Your knowledge hub

The European Digital Single Market

" The digital economy revolves around data. It is the driving force behind those three main elements of productivity, innovation and digitalisation.

Let's not lose time being afraid - let's build an open and vibrant data economy."

> Understanding the economic impact of data

> Inclusive Economic Growth and Job Creation

See: I https://ec.europa.eu/commission/2014-2019/ansip/announcements/speech-vice-presi- dent-ansip-bruegel-annual-meeting-productivity-innovation-and-digitalisation-which_en

Creating legal frames

DIRECT IMPACT OF LEGAL FRAMES:

This legal frame contributes to the creation of:

- a common legal space for *FAIR data* and *FAIR infrastructures*

- *crossborder services* and *related infrastructures*

- the use of *common terminologies*

- the *alignement of policies*

- the adoption of *good governance models* along the whole data value chain

- the adoption of *funding mechanism* that lead to the distribution of data

- the realization of goals pursued by the *Open Access* movement

The European Digital Single Market Added Value for All

Compared to other regions across the globe, the European Union (EU) Member States are generally characterized by an above-average participation of the public sector in the national economies, with common estimates of the public sector's stake in national GDPs ranging from 25% to 50% of their respective economies, depending on the country and on the metrics used.

As a result, European administrations generally invest a significant budget in the creation of **Public Sector Information (PSI),** or **Open Data**.

> *This information has a potential economic value that significantly exceeds its strict public sector utility.*

Source: EOSC-SRIA V1.0, p. 150

In 2003, the EU issued legislation to govern the publication of Open Government Data in Member States via the so-called Public Sector Information (PSI) Directive 2003/98/EC.17 The main objective was to enable better access to Open Data by:

Stimulating the further development of a European Market for Open Data based services

Enhancing the cross-border use and application of Open Data in business processes

Encouraging competition in the internal market

Addressing the differences in rules and practices between Member States

Creating legal frames

The **DIRECTIVE EU 2013/37** *amending Directive 2003/98/EC on the re-use of public sector information*

This initiative contains crucial elements that are relevant. Some of them contribute to support the creation of

- a legal frame for the creation and the distribution of data (→ open data),

- the creation of related services (→ distributive by design),

- infrastructures (→ accessible infrastructures for FAIR use)

The Directive EU 2013/37

Among all legal instruments, the Directive EU 2013/37 - which amended the Directive 2003/98/EC on the re-use of public sector information has demonstrated to one of the most crucial instruments

According to paragraph (33) the objectives of Directive EU 2013/37 are to

- facilitate the creation of Union-wide information products and services based on public sector documents, and
- to ensure the effective cross-border use of public sector documents

Source: EOSC-SRIA V1.0, p. 150

Furthermore the Directive explicitly recognizes the role of some major **stakeholder** in the processes of data creation, management and re-use.

The Directive does not only refer directly to the data, their use and maintenance, but it foresees the **automation of mechanisms** for the maintenance and the accessibility of data and it includes also the possibility of **PPP models** and resulting business models.

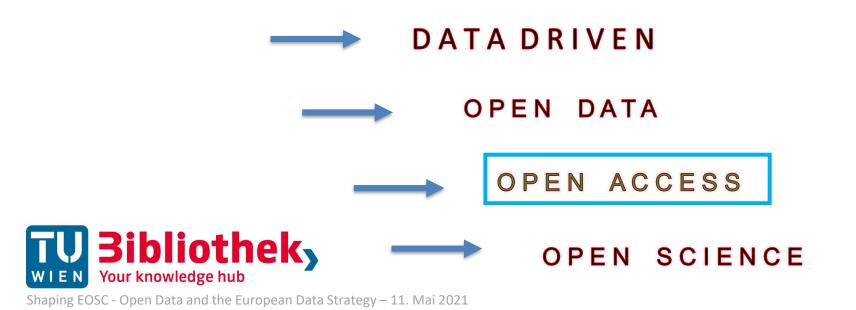*Source: EOSC-SRIA V1.0, p. 150*

Aligned with the EOSC Objectives

→ Build resilient infrastructure,

→ promote sustainable development

→ fostering innovation in e-infrastructure

→ e-Infrastructures, especially research infrastructures, **findable, accessible, interoperable and reusable** across national and disciplinary boundaries.

DATA DRIVEN

OPEN DATA

OPEN ACCESS

OPEN SCIENCE

Among all stakeholder, the Directive assigns → a relevant to University Libraries

(University) libraries, museums and archives are mentioned eight times in the directive itself and five times in the amendment of the Directive

University Libraries are relevant stakeholders in the Digital Single Market and in the European Open Science Cloud

They also play a relevant role in the Open Science mouvement because they hold the data

They know how to manage them, to enrich them

They can assume roles of brokerage of data and of services

Libraries are the main stakeholders of huge pan-European e-Infrastructures Projects like OpenAIRE, RECODE, PASTEUR4OA, LEARN or are represented in many initiatives led by CESAER LERU or LIBER
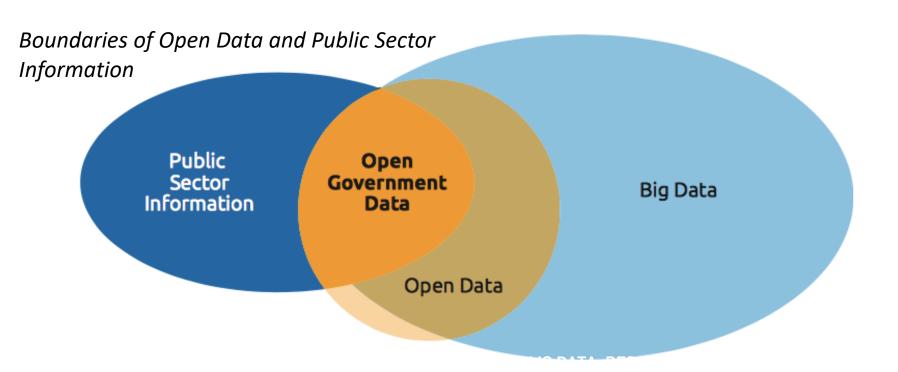
The added value

Creating Value Through Open Data

Open Data is a sub-set of the commonly used term Big Data

Boundaries of Open Data and Public Sector Information

Public Sector Information

Open Government Data

Open Data

Big Data

EOSC – Creating value though Open Data

European Science Cloud is a way to achieve higher level of economic productivity through technology and innovation, in areas of like i.e.Trade, Tourism, Financial Services, Manufactoring, ICT..., and on drafting development oriented strategies that support innovation, creativity and job creation.

Though the creation and the management of data and related services may be expensive, Open Science, Open Data and the EOSC will pave the way to substantial economic growth, as countries like France the U.S. and others are experiencing. See e.g. the study *Creating Value through Open Data* from the European Data Portal, created by the DG Connect of the EU Commission, available at this website: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

The value generated by Open Science and the EOSC through Open Data

Several benefits of the re-use of Open Data and related services are identified in the study of the Commission. These consist of direct and indirect benefits.

Direct benefits are monetised benefits that are realised in market transactions in the form of revenues and Gross Value Added (GVA), the number of jobs involved in producing a service or product, and cost savings.

Indirect economic benefits are i.e. new goods and services, time savings for users of applications using Open Data, knowledge economy growth, increased efficiency in public services and growth of related markets:
more transparency in the decision making processes

Open Data Portals

The first country to have an Open Data portal was Spain in 2009. They are also one of the frontrunners in terms of their Open Data policy.

2009	2010	2011	2012	2013	2014	2015	≥2016
Spain	Slovenia	Belgium		Austria	Bulgaria	Croatia	Iceland
	UK	Estonia		Denmark	Cyprus	Czech Republic	Latvia
		France		Germany	Finland	Hungary	Liechtenstein
		Italy		Greece	Ireland	Lithuania	Luxembourg
		Netherlands		Romania	Poland		Malta
		Norway		Slovakia			
		Portugal		Sweden			
				Switzerland			

Creating value through Open Data

For 2016 the total market value of Open Data was estimated between 193 bn EUR and 209 bn EUR with an estimated projection of 265-286 bn EUR for 2020, including inflation corrections.

For the period 2016-2020, the cumulative direct market size was estimated at 325 bn EUR. The cumulative total market size for Open Data is forecasted to be between 1,138 and 1,229 bn EUR.

Source: **November 2015** - A study on the Impact of Re-use of Public Data Resources
https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

€ 325 billion direct market size EU28+ for 2016-2020

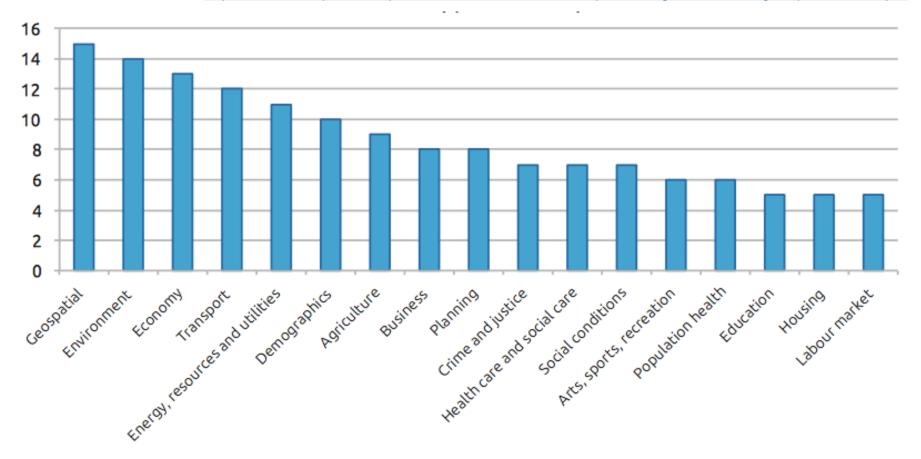# The value generated by Open Data

High

Commerical Re-use of Open Data

Geographic information
Meteorological and Environmental information
Economic and Business information
Social information
Traffic and Transport information
Tourist and Leisure information
Agricultural, Farming, Forestry and Fisheries information
Natural Resource information
Legal System information
Scientific information and Research data
Educational content
Political content
Cultural content

→ *Open Data: relevant reuse divided by disciplines and sectors*
→ *Data coming out of publicly funded research are to be considered PSI*
→ *Foster **Open Access***

Sectors most applicable to Open Data

Numbers of sectors to which the data is applicable as identified by Deloitte , (Deloitte 2013)

Open Data – Environmental Benefits

Open Data help achieve environmental benefits

It is shown that Open Data has the potential of saving 1,425 lives a year (i.e. 5,5% of the European road fatalities).

Furthermore, applying Open Data in traffic can save 629 million hours of unnecessary waiting time on the road in the EU.

Source: **November 2015** - A study on the Impact of Re-use of Public Data Resources
https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
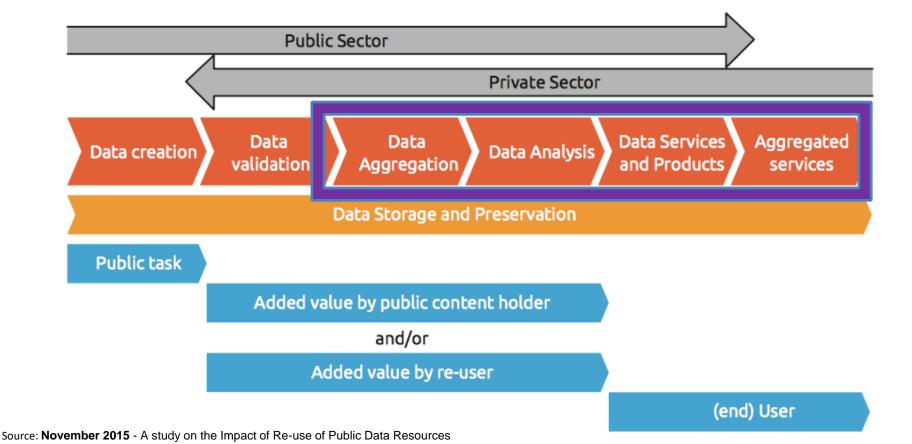
Making Use of Open Data

The Data Value Chain can serve as a basis to understand different types of re-use

Open Data Value Chain

CapgeminiConsulting,basedon:MEPSIR(2006),p.46andhttp://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=3488

Data creation is done by the *Suppliers*. The organisations that collect and aggregate the Open Data are called the *Aggregators*. Individuals or companies that analyse the data and create services and products can be divided into *Developers* and *Enrichers*

Data Value Chain Archetypes
http://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf

THANKS!

Paolo Budroni
TU Wien Bibliothek

paolo.budroni@tuwien.ac.at

Head of EOSC and International Liaison Office

Austrian e-IRG Delegate – e-IRG Chair

https://orcid.org/0000-0001-7490-5716

TU WIEN Bibliothek
Your knowledge hub