Architectures of Knowlege
Chapter II – Sustainable Growth

The Data Way to Open Science and the Economic Impact of Open Data

Paolo Budroni
Vienna University Library
Dpt. e-Infrastructures

Abstract – Architectures of Knowlege
Chapter II – Sustainable Growth
The Data Way to Open Science and the Economic Impact of Open Data

The European Commission is promoting the Open Science, Open Data and the European Open Science Cloud. These will be soon a reality. 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 will be 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 and understand the economic impact of Open Data. This presentation aims at analysing these settings, first embedding them in a broader international legal and economic context related to the implementation of policies, and then offering some insights and figures concerning the value of an Open Data driven European_economy.

Why this slides?

Understanding Open Data, The European Digital Single Market, Open Science, the European Open Science Cloud and the value of an Open Data driven European economy.

The European Digital Single Market

Andrus Ansip,
Vice-President Digital Single Market

"If I had to express my views about the digital future – that of Europe or indeed, of the whole world - I could do it with one word: data."

" 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."

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

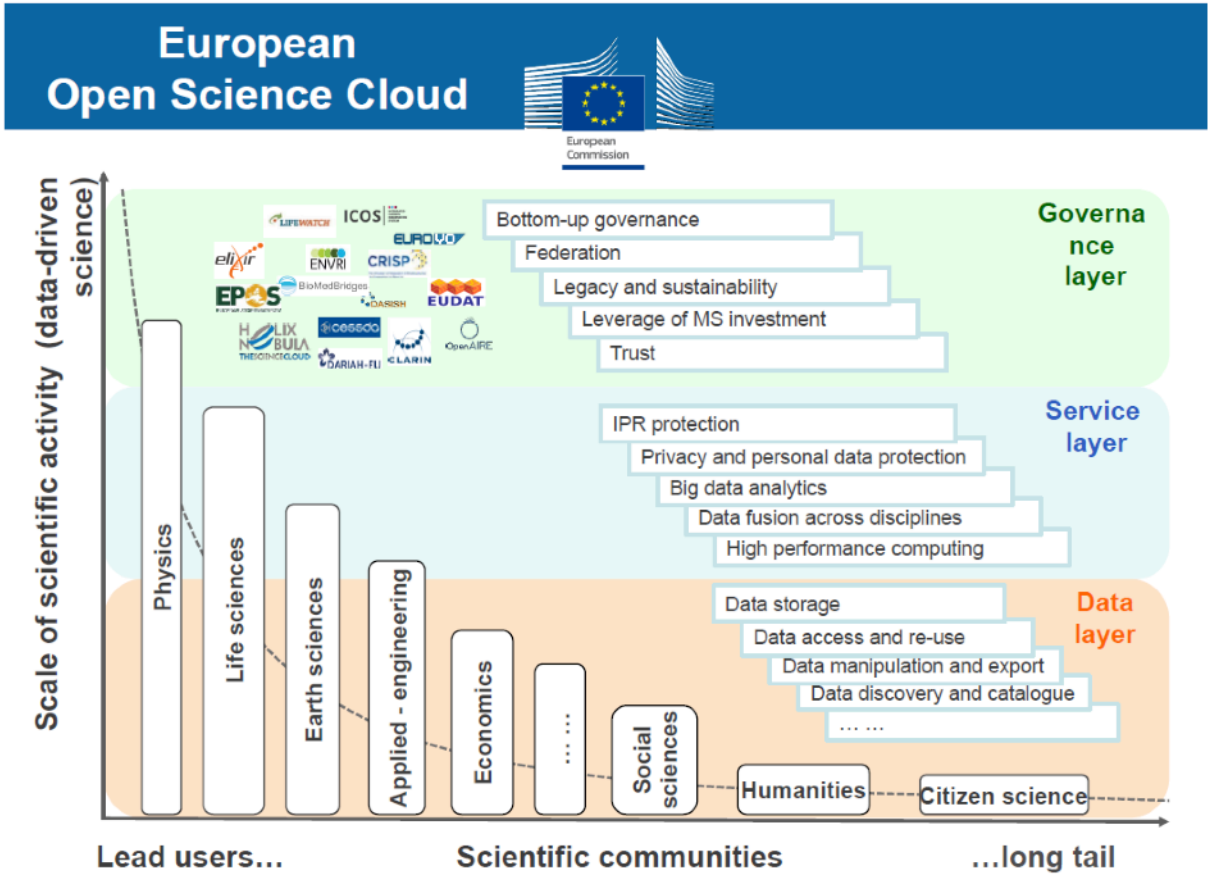
Open Science and the European Open Science Cloud

The European Commission is promoting the European Open Science Cloud. This will be soon a reality. 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 will be 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

European Open Science Cloud



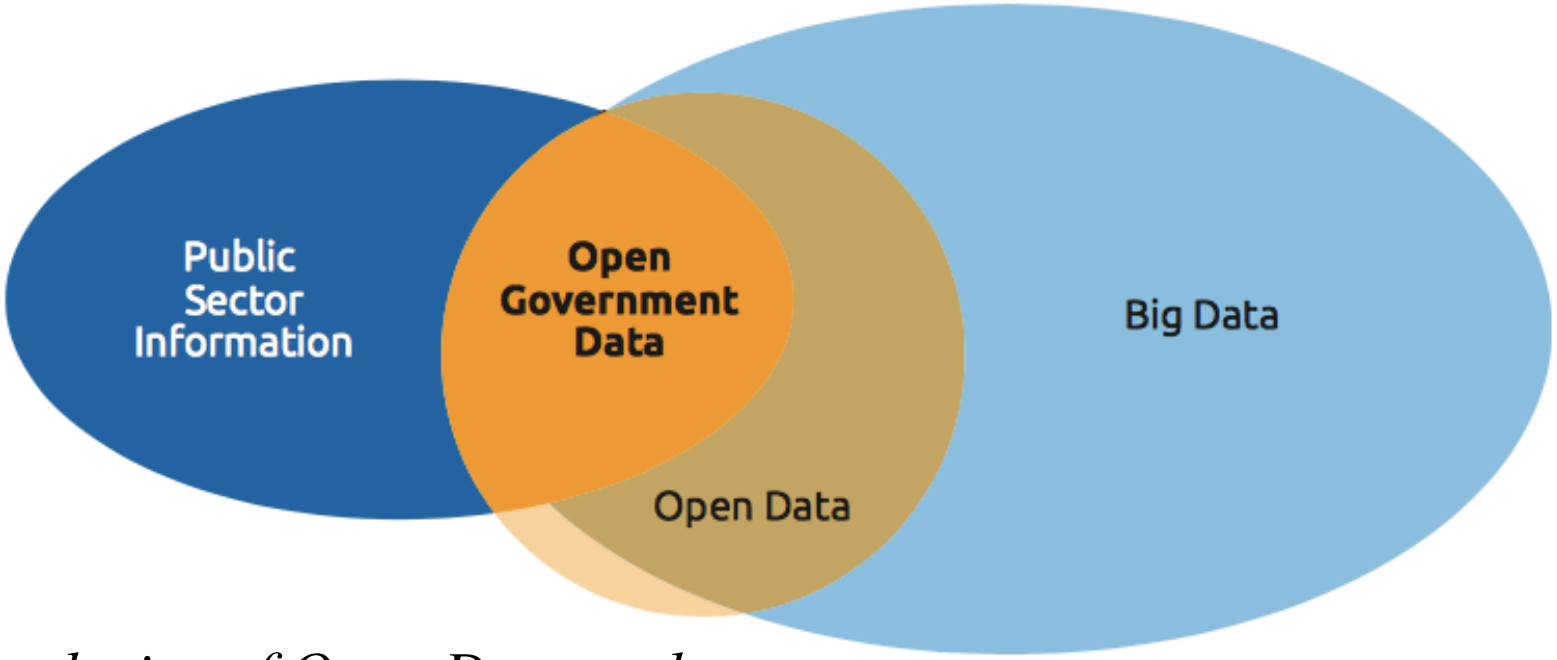
Source: "Open Science policy: Results of the consultation on 'Science 2.0: Science in transition' and possible follow up." Presented by J.C. Burgelman, June 3 2015 at e-IRG workshop.

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.2 As a result, European administrations gener- ally invest a signi cant budget in the crea- tion of Public Sector Information (PSI), or Open Data. This information has a potential economic value that signi cantly exceeds its strict public sector utility.

Creating Value Through Open Data

Open Data is a sub-set of the commonly used term Big Data
→ Big Data, Open Data and ODGs are also produced at Research Institutions



Boundaries of Open Data and Public Sector Information

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

Open Data – Impact on Job Creation

EU 28+ private sector: job creation

→ directly linked to the re-use of Open Data - excluding staff working in government administrations

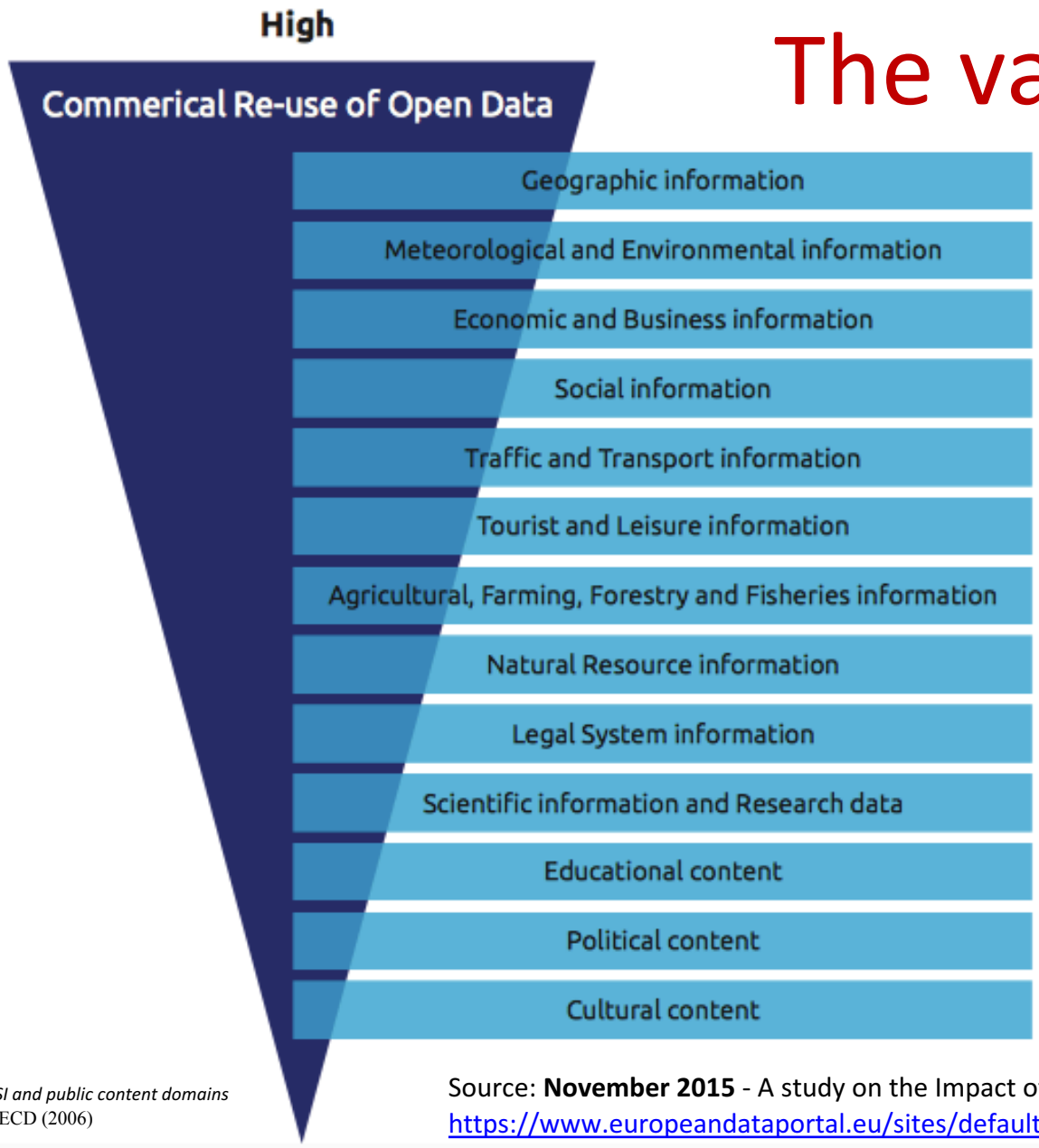
In 2016 The forecasted total number of direct Open Data jobs **in 2016** has an upper bound of 75,000 jobs.

By 2020, the upper bound provides a forecast of just under 100,000 direct jobs directly related to Open Data jobs. Thus, in the period 2016-2020, almost 25,000 direct Open Data jobs are created.

This equals a growth of 32% over a 5-year period. Per year, the increase in number of jobs is forecasted to grow at an average rate of 7.3%.

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

The value generated by Open Data

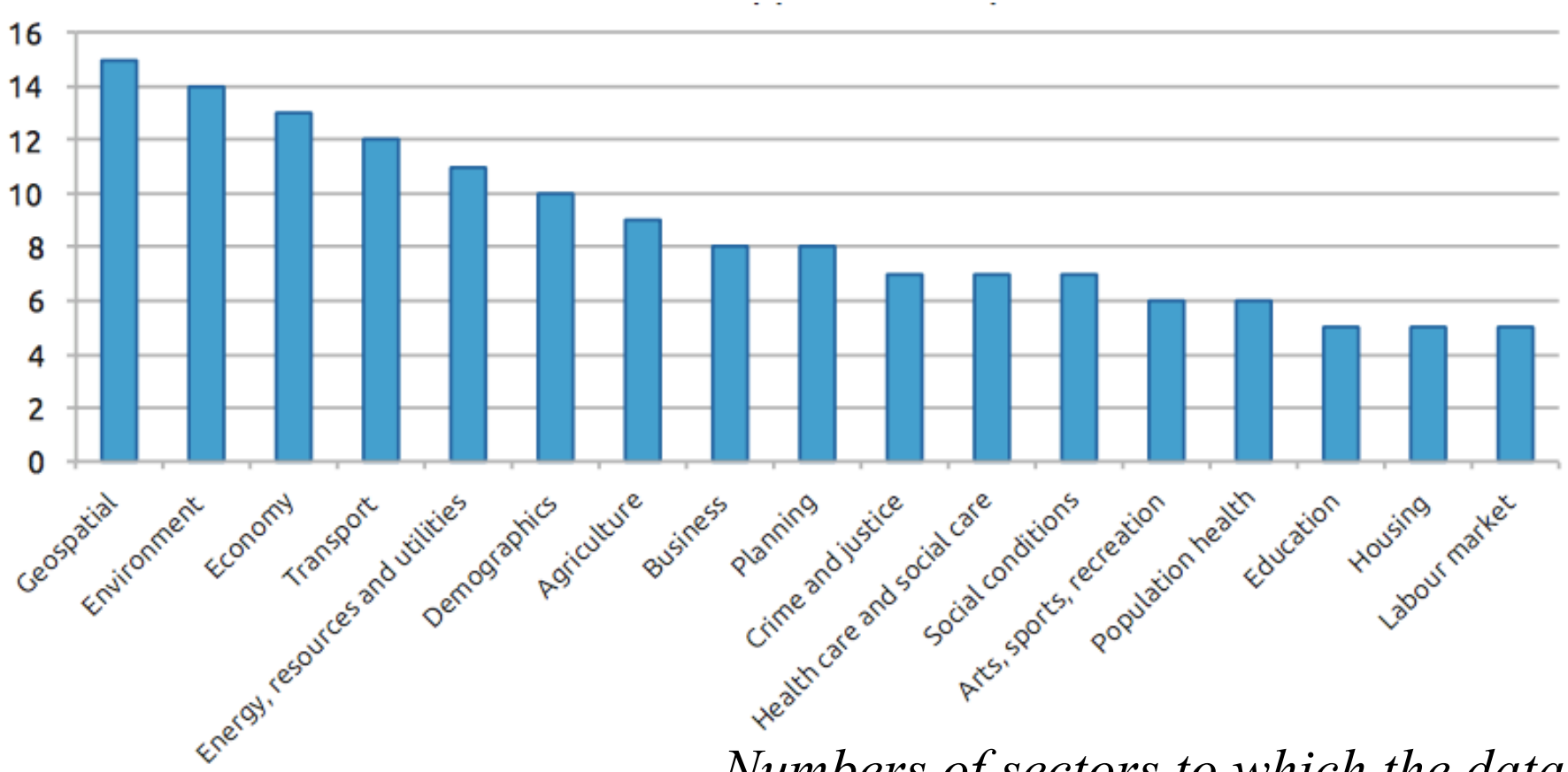


PSI and public content domains
OECD (2006)

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

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

Sectors most applicable to Open Data

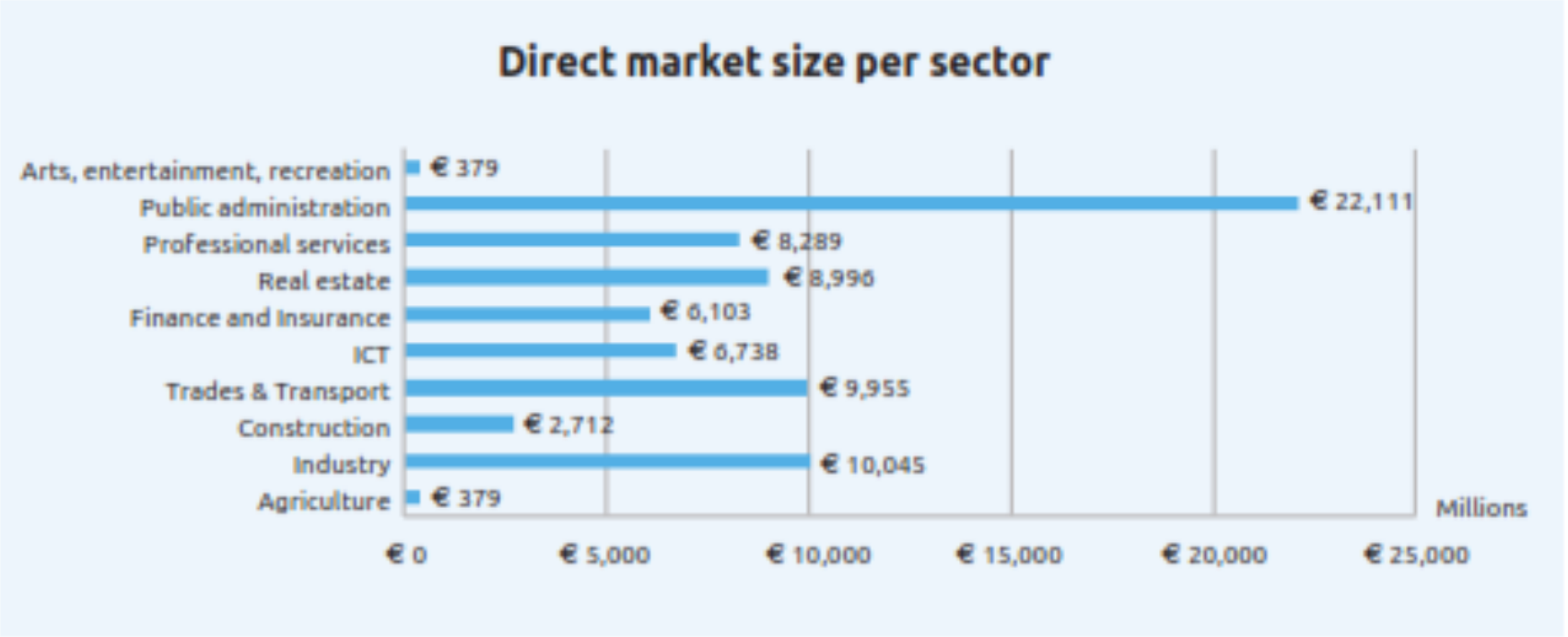


Numbers of sectors to which the data is applicable as identified by Deloitte , (Deloitte 2013)

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

The value generated by Open Data

Between 2016 and 2020 25,000 jobs directly related to Open Data will be created



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

Open Science, Open Data and EOSC: Impact on the economic world

The 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

The European Data Portals

Within the context of the launch of the European Data Portal, further evidence of the quantitative impact of re-use of Open Data was measured. The aim of the study was to collect, assess and aggregate economic evidence to forecast the benefits of the re-use of Open Data for the EU28+. Four key indicators were measured:

- direct market size,
- number of jobs created,
- cost savings,
- efficiency gains.

Between 2016 and 2020, the market size of Open Data is expected to increase by 36.9%, to a value of 75.7 bn EUR in 2020. The forecasted number of direct Open Data jobs in 2016 is 75,000 jobs. From 2016 to 2020, almost 25,000 extra direct Open Data jobs are created. The forecasted public sector cost savings for the EU28+ in 2020 are 1.7 bn EUR. Efficiency gains are measured in a qualitative approach.

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

Creating Value Through Open Data

For 2016 the total market value of Open Data is 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 is 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

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

The value generated by Open Data

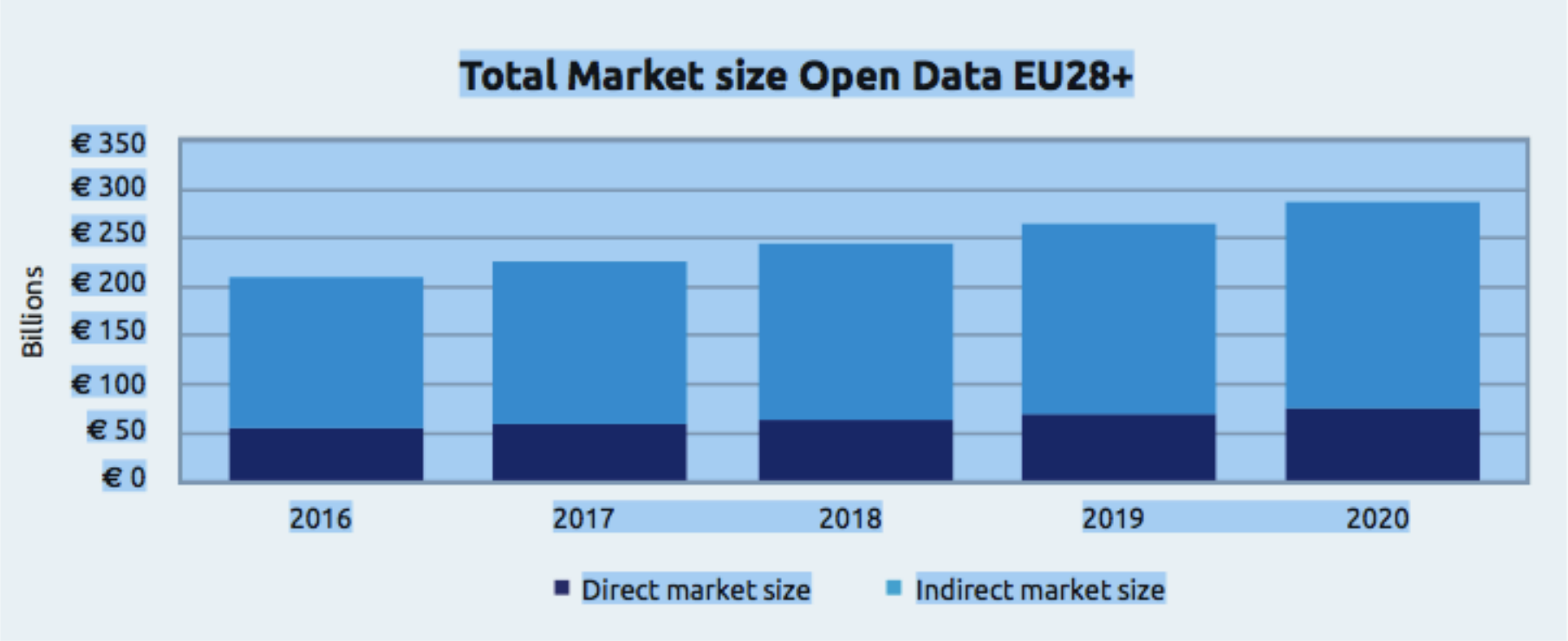
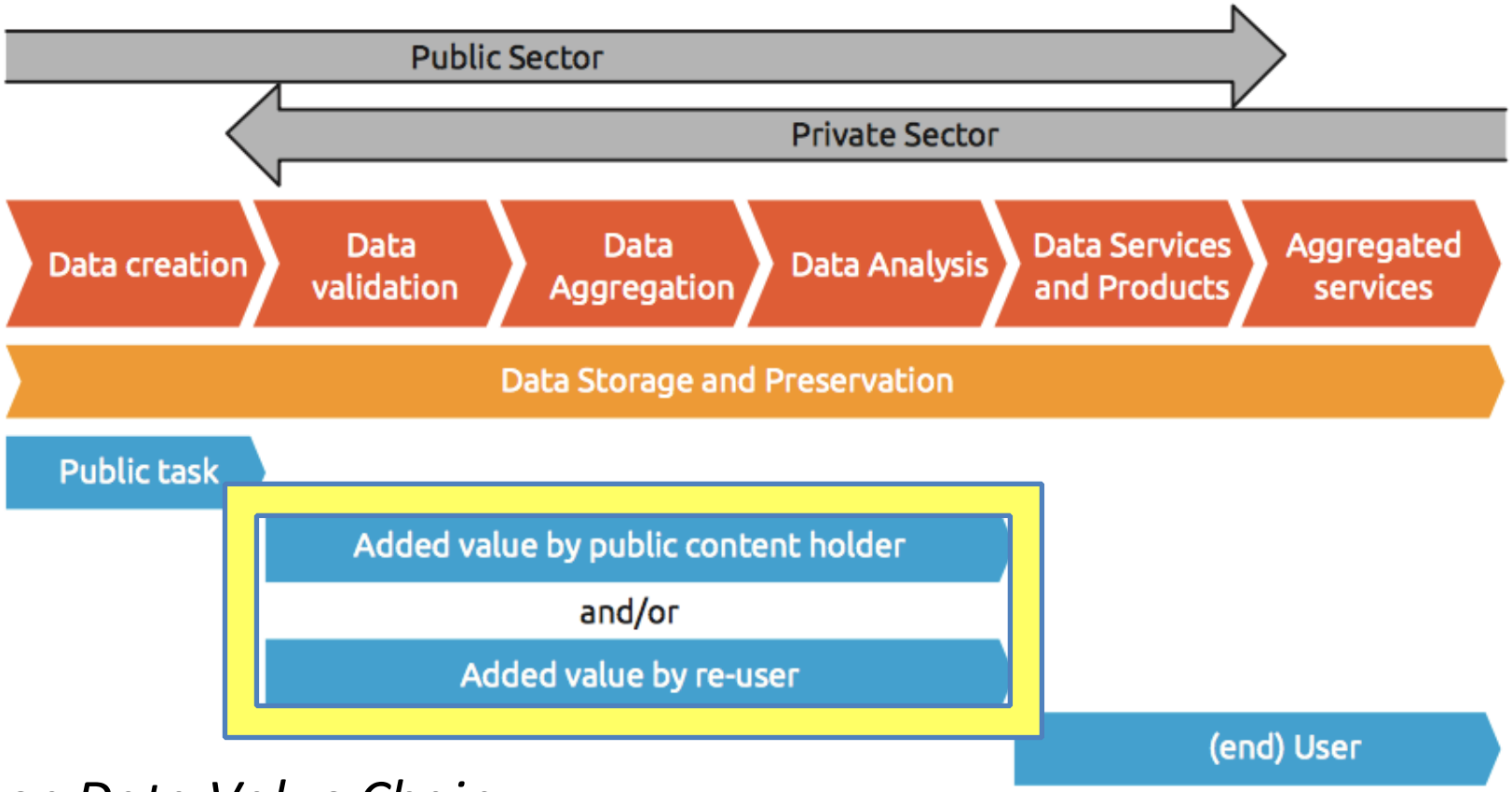


Figure 1 – Total market size (high bound), split in direct and indirect size for EU28+in billions, 2016-2020
Figure 1 - Marché total (fourchette haute), direct et indirect pour l'UE28+ en milliards, 2016-2020

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

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

The Data Value Chain

The creation of data services and products

Three forms of Open Data can be distinguished and are related to the Data Value chain

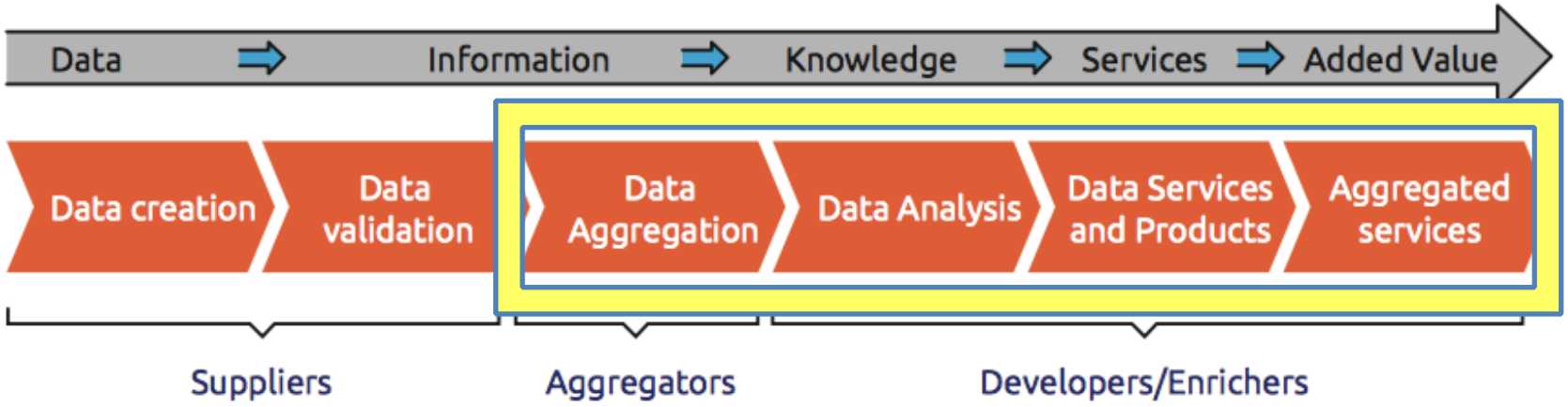
- Raw data corresponds to quantities or other quantitative or qualitative attributes derived from observation, experiment, measurement or computation. The data are not structured, contextualised nor commented.
- Information corresponds to a set of contextualised and structured data, the producer's intention being to make them meaningful.
- Knowledge corresponds to cognitive appropriation of the information by an individual who organises, synthesises and/or summarises it to make it more readily understandable.

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

The Data Value Chain

The creation of data services and products

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

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

Creating Legal Frameworks

On a high level two initiatives with related legal aspects have a huge impact on the architectures proposed by the EOSC initiative and the Open Science movement

1. The **SDG** (Sustainable Development Goals) of the U.N.

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

Creating Legal Frameworks

These **1+2 initiatives** contain 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), and of course infrastructures (→ accessible infrastructures for fair use)

DIRECT IMPACT:

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

SUSTAINABLE DEVELOPMENT GOALS
17 GOALS TO TRANSFORM OUR WORLD

Home About Secretary-General Goals Take Action News and Media Watch and Listen

1

SUSTAINABLE DEVELOPMENT GOALS

tainabledevelopment/hunger/

→ **Three SDG are relevant in this context: #9 #4 #5**

Industry Innovation & Infrastructure

- Quality infrastructure is positively related to the achievement of social, economic and political goals
- Inadequate infrastructure leads to a lack of access to markets, jobs, information and training, creating a major barrier to doing business
- Undeveloped infrastructures limits access to health care and education

Quality Education & Gender Equality

- By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university
- By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situation

1

Sustainable Development Goals (SDG) and EU-COM

1

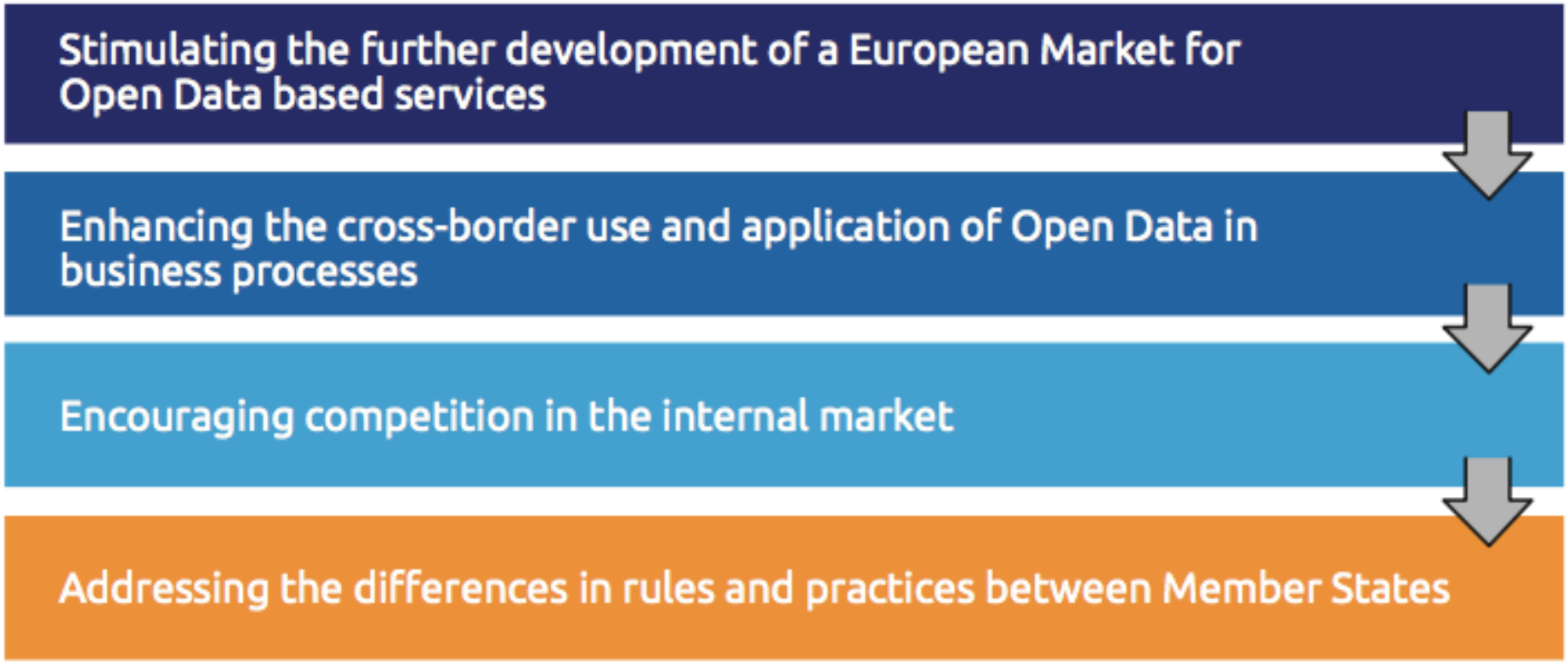
Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

Blending, which combines EU grants with loans or other public and private funding, is increasingly supportive of actions on infrastructure which have a multiplier effect on sustainable development in partner countries.
EU actions are geared towards inclusive and sustainable growth and economic integration, building on partner countries' comparative advantages in the manufacturing or services sectors. Other targets on access to financial services and technology are consistent with existing EU policy, including its involvement in initiatives such as the Global Partnership for Financial Inclusion.

→ AND IT IS ALL BASED ON DATA

Creating Legal Frameworks

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:



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

Creating Legal Frameworks

„The European PSI Directive (2003) was established on a minimum harmonisation basis allowing Member States to pave the way to the **interpretation and implementation of the framework**. The European Directive established framework rules regarding availability, accessibility and transparency of Open Data in Europe. In addition, it was recommended to have a standard electronic licence for the re-use of Open Data and to have a tool to access the relevant data sets via a list of portal websites.„

→ in 2017: it is required that Data coming out of publicly funded research are to be considered Public Sector Information

2

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

The DIRECTIVE EU 2013/37

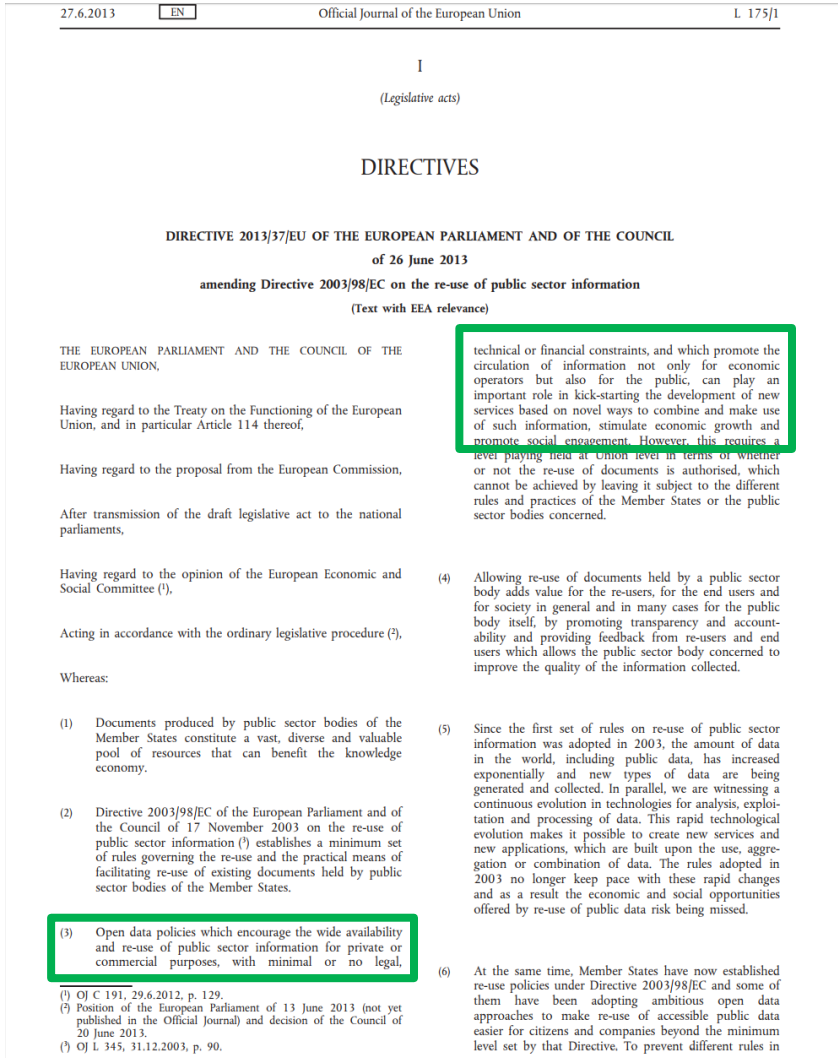
amending Directive 2003/98/EC on the re-use of public sector information

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 be a crucial instrument

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

(§3) Open Data Policies



27.6.2013 EN Official Journal of the European Union L 175/1

I

(Legislative acts)

DIRECTIVES

DIRECTIVE 2013/37/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
of 26 June 2013
amending Directive 2003/98/EC on the re-use of public sector information
(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 114 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee (1),

Acting in accordance with the ordinary legislative procedure (2),

Whereas:

(1) Documents produced by public sector bodies of the Member States constitute a vast, diverse and valuable pool of resources that can benefit the knowledge economy.

(2) Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information (3) establishes a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public sector bodies of the Member States.

(3) Open data policies which encourage the wide availability and re-use of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints, and which promote the circulation of information not only for economic operators but also for the public, can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulate economic growth and promote social engagement. However, this requires a level playing field at Union level in terms of whether or not the re-use of documents is authorised, which cannot be achieved by leaving it subject to the different rules and practices of the Member States or the public sector bodies concerned.

(4) Allowing re-use of documents held by a public sector body adds value for the re-users, for the end users and for society in general and in many cases for the public body itself, by promoting transparency and accountability and providing feedback from re-users and end users which allows the public sector body concerned to improve the quality of the information collected.

(5) Since the first set of rules on re-use of public sector information was adopted in 2003, the amount of data in the world, including public data, has increased exponentially and new types of data are being generated and collected. In parallel, we are witnessing a continuous evolution in technologies for analysis, exploitation and processing of data. This rapid technological evolution makes it possible to create new services and new applications, which are built upon the use, aggregation or combination of data. The rules adopted in 2003 no longer keep pace with these rapid changes and as a result the economic and social opportunities offered by re-use of public data risk being missed.

(6) At the same time, Member States have now established re-use policies under Directive 2003/98/EC and some of them have been adopting ambitious open data approaches to make re-use of accessible public data easier for citizens and companies beyond the minimum level set by that Directive. To prevent different rules in

(1) OJ C 191, 29.6.2012, p. 129.
(2) Position of the European Parliament of 13 June 2013 (not yet published in the Official Journal) and decision of the Council of 20 June 2013.
(3) OJ L 345, 31.12.2003, p. 90.

→ The need of

"Open Data Policies which encourage the wide availability of re-use of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints, and which promote the circulation of information not only for economic operators but also for the public…"

(§3) Open Data Policies *and economic growth*

(3) Open data policies which encourage the wide availability and re-use of public sector information for private or

l,

technical or financial constraints, and which promote the circulation of information not only for economic operators but also for the public, can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulate economic growth and promote social engagement. However, this requires a

2

(§17) Harmonisation of rules and practices →*policies*

(17) Since the differences in national rules and practices or the absence of clarity hinder the smooth functioning of the internal market and the proper development of the information society in the Union, minimum harmonisation of national rules and practices on the re-use of public cultural material in libraries, museums and archives should be undertaken.

2

The DIRECTIVE EU 2013/37

amending Directive 2003/98/EC on the re-use of public sector information

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.

(§18) Scope II → *new roles for libraries*

(18) The extension of the scope of Directive 2003/98/EC should be limited to three types of cultural establishments – libraries, including university libraries, museums and archives, because their collections are and will increasingly become a valuable material for re-use in many products such as mobile applications. Other

Among all stakeholder, the Directive assigns →
a relevant to University Libraries, quoted 8 times

2

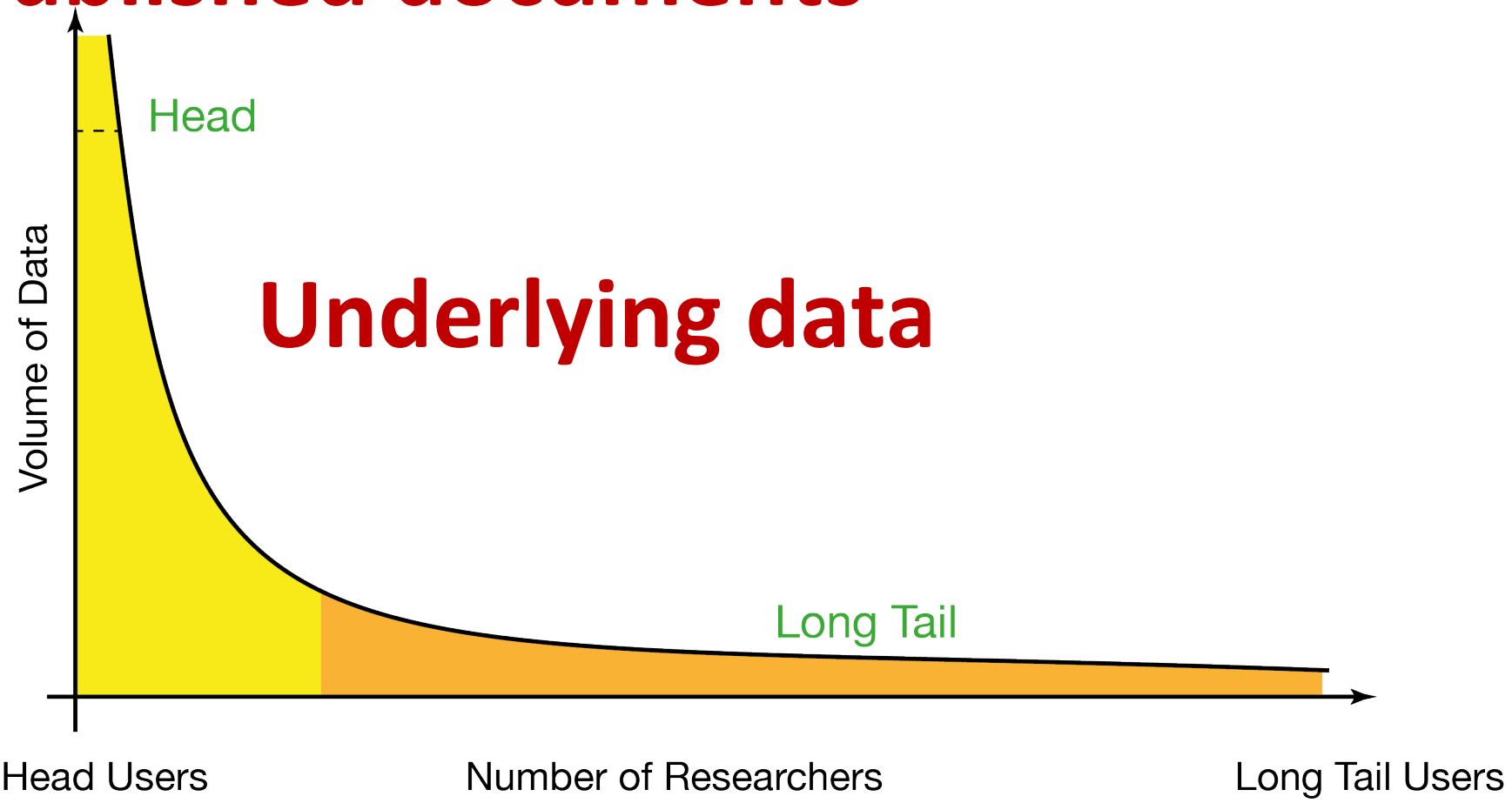
Among all stakeholder, the Directive assigns → a relevant to University Libraries,

- *(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 stgakeholders of huge pan-European e-Infrastructures Projects like OpenAIRE, RECODE, PASTEUR4OA, LEARN or are represented in many initiatives led by LERU or LIBER*

2

Open Data / Research Data
Libraries as Holders of Data

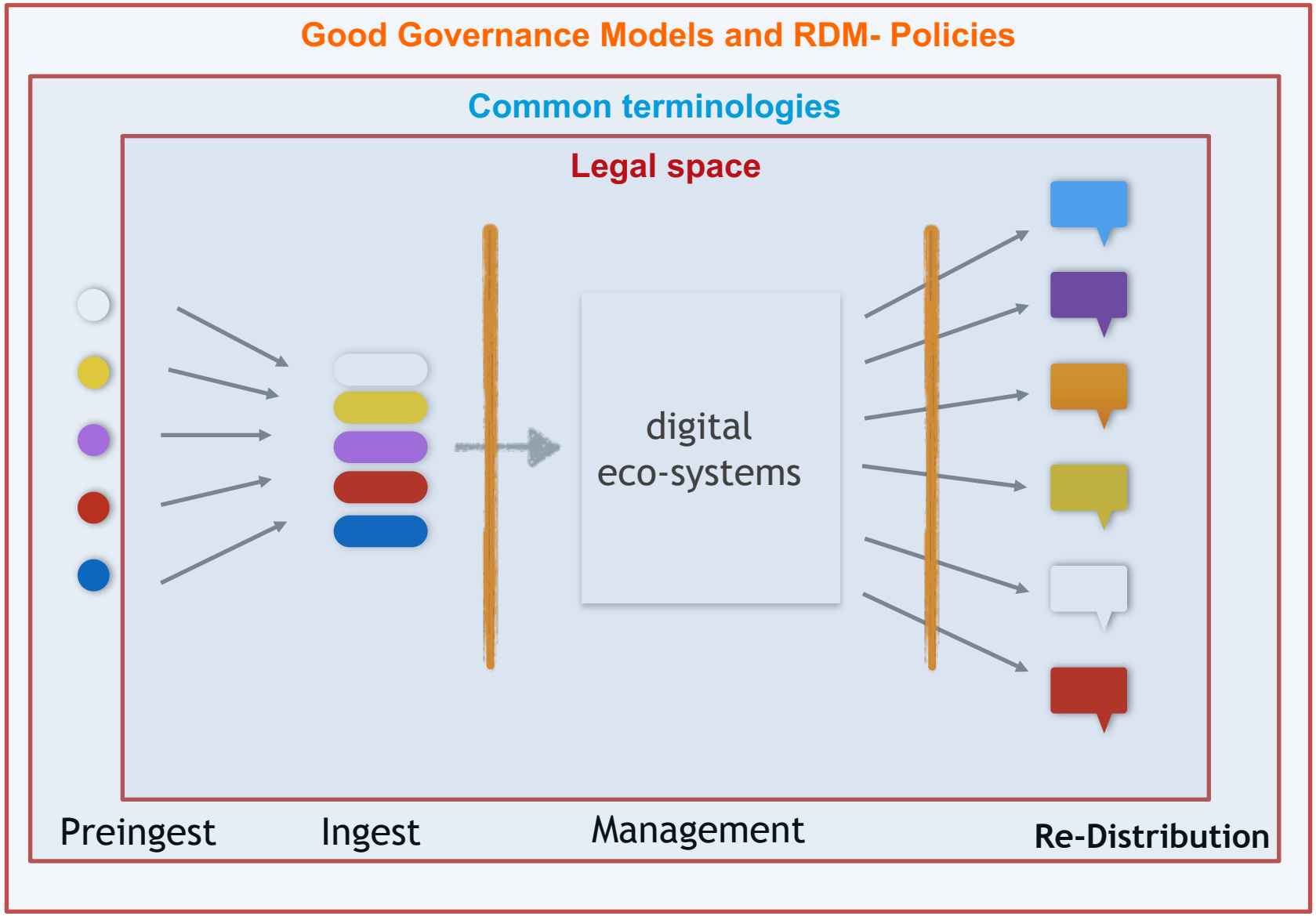
Published documents



Sustainable Growth: Conclusions

Needed:
Assignments of new roles, responsibilities and duties

RDM-Policies



Towards Sustainable Growth - Steps enhancing a strategy development

- Identify small and medium players for the upcoming EOSC
- Understand good governance models and implement them
- Generate and adopt Key Performance Indicators (KPIs)
- Offer business a matrix against which to score their sustainability
- Concerning business scorecards: push corporate ambition in the right direction in terms of measuring what matters
- Create a network/ a structure of flows
- Restructure: create an economy that is distributive by design, embedded in a web of flows
- Generate a common legal space for services and for data
- Enhance permanent access
- Introduce the FAIR principles
- Enhance accessibility
- Involve Data Support Units: ICT Services, **Libraries**

Ten Recommendations for Strategy and Policy Development

- Understand Open Science on a broad scale
- Become „Partner of Reference“ for the EOSC
- Concerning “Citizen Science” → empower qualified people
- Develop skills for lateral technologies and achieve distributive design
- Teach social entrepreneurship
- Equip young people with these skills
- Plan for re-distribution of data
- Create for regenerate
- Create a digital agenda and an Open Digitization Policy
- Ensure that all publicly funded research becomes public knowledge – introduce public domain mark (CC0 1.0) wherever it is possible

A presentation produced by the Austrian Open Science Support Group (AOSSG) Coordination Team:

Paolo Budroni, University of Vienna, Library and Archive Services
Head of Department e-Infrastructures, member of coordination team AOSSG
Austrian National Delegate of E-IRG (E-Infrastructures Reflection Group)
Chair of GA of e-Infrastructures Austria Plus, H2020 Project LEARN – WP3 Policy Development and Alignment, Member of Management Board of Open Education Austria, TAIEX Expert

Barbara Sánchez Solís, University of Vienna, Library and Archive Services, Manager Austrian Social Science Data Archive (AUSSDA), member of coordination team AOSSG, former coordinator of e-Infrastructures Austria, leader ov WP2 e-Infrastructures Austria Plus, Data Management Plans (DMP)

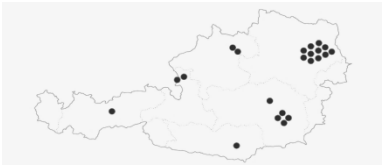
Raman Ganguly, Head of Design and Software Dpt. at Vienna University Computer Centre, member of coordination team AOSSG
Leader of WP3 Open Education Austria, former Technical Director of e-Infrastructures Austria

AOSSG – Aknowledgement and Contacts

Responsible for the content of this version (version November 2017):

Paolo Budroni, Vienna University, Library and Archive Services
in cooperation with
Raman Ganguly, Vienna University, Computer Centre
and
Barbara Sánchez Solís, Vienna University, Library and Archive Services

Paolo Budroni, paolo.budroni@univie.ac.at
Raman Ganguly, raman.ganguly@univie.ac.at
Barbara Sánchez Solís barbara.sanchez.solis@univie.ac.at



Authors of the graphics:
Paolo Budroni
Raman Ganguly

Credits to Melanie Stummvoll, Vienna University Library and Archive Service, for the participation at the realisation of the slides concerning the EU Directive 2013/37

THANKS!

Paolo Budroni, paolo.budroni@univie.ac.at

University of Vienna, Library and Archive Services
Head of Department e-Infrastructures
E-Infrastructures Reflection Group, Austrian National Delegate

Chair of GA of e-Infrastructures Austria
H2020 Project LEARN – WP3 Policy Development and Alignment
Member of Management Board of Open Education Austria
TAIEX Expert