

A background network diagram consisting of several grey circular nodes of varying sizes connected by thin grey lines. The nodes are arranged in a non-uniform pattern, with some larger nodes and some smaller ones, creating a web-like structure.

Enabling Open Science

The Data Deluge: the Role of Research Organisations

Paolo Budroni - Caribbean Research Data
Workshop on LEARN Project 24/11/2016

e-infrastructures
austria



Credits:

This presentation contains some three slides (5, 6, 30) which were created by Dr. Paul Ayris, UCL. Paul Ayris is member of the *High Level Expert Group on European Open Science Cloud* and PI of the *Project LEARN*.

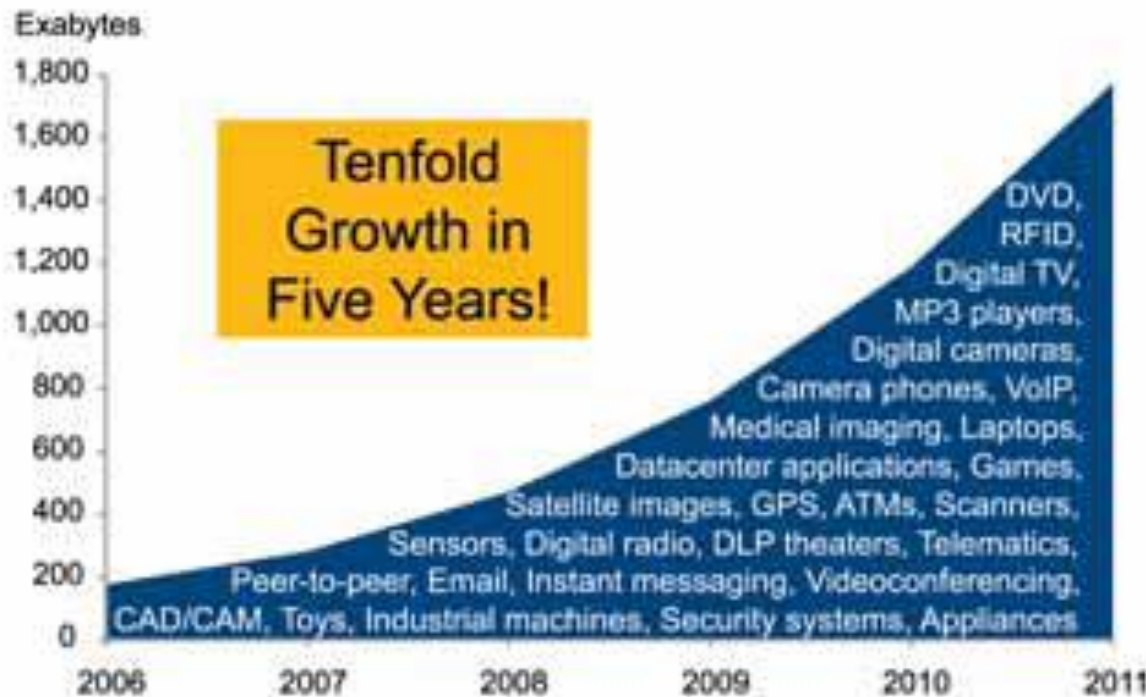
These 3 slides were also part of his presentation at the **4th LEARN Workshop** in Santiago de Chile - 27th October 2016

1. Data deluge and research data
2. A case: Austrian Open Science Infrastructure - What is really needed?
3. About Open Science: The European Open Science Cloud
4. Findings and conclusions. About Policies,
the Project LEARN

Data deluge and research data



Digital Information Created, Captured, Replicated Worldwide



The internet has 1800 exabytes of data in 2011

exa = 10^{18}

Steven Vale

UNECE

steven.vale@unece.org

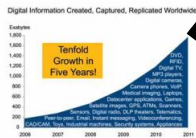


UN Economic Commission for Europe

50,000 exabytes by 2020

**27 fold
growth in
the next
9 years**

**We live in
exponential
times!**



Different levels of processing of data Model for digital archiving

World of data

Raw data (primary data)

Processed Data
Inconclusive
Results

Processed Data

Processed Data
Negative Results

Positive results

Positive results

Released
Data

Shared
Data

Pub.
Data

Shared
Data

Pub.
Data

Shared
Data

→ Strata of research data

→ Restricted Data

→ Open data

→ Published data

→ Open access published data

OA

Ensuring legal and ethical compliance is key issue in this context



INVOLVED STAKEHOLDER

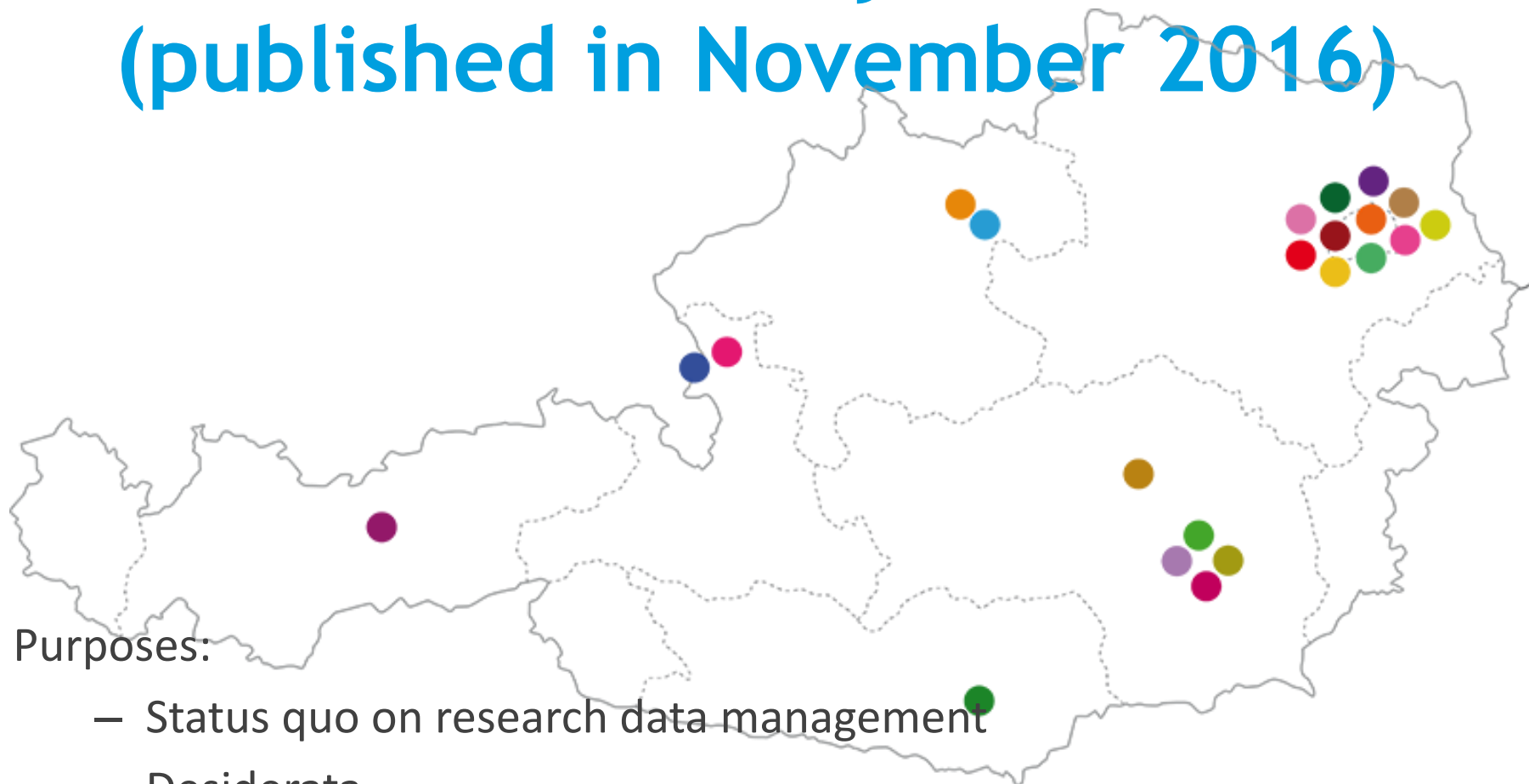
- **Universities**
- **Non-university research institutions**
- **Research Communities**
- **Research-funding bodies**
- **Research Support Institutions or Services**
 - Scientific Libraries, IT-Services, Research Support Services, Legal Services

**A case:
Austrian Open Science
Infrastructure - What is really
needed?**



Austrian National Research Data Survey

(published in November 2016)



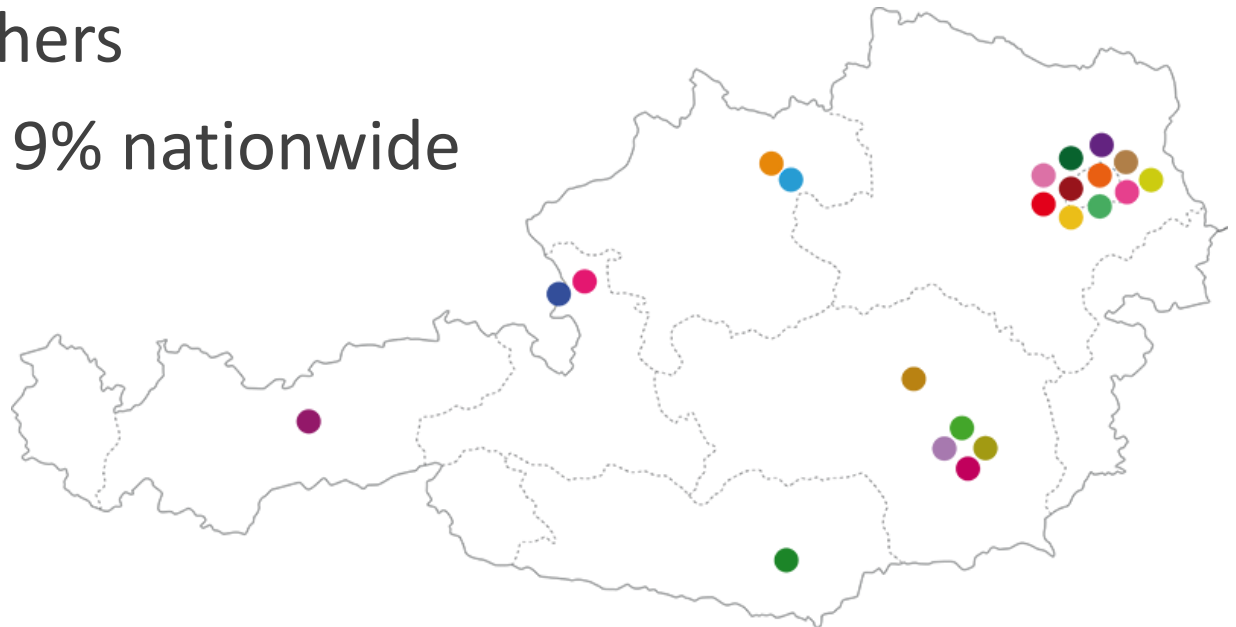
Purposes:

- Status quo on research data management
- Desiderata
- Enabling open science
- Raise awareness

e-infrastructures
austria

Target and response

- Researches of all 21 public universities and three extra-university research institutions in Austria
- From post doc level upwards
- According to knowledge management report: 36.000 researchers
- Response rate: 9% nationwide



Austrian National Survey published in November 2015

- Purposes:
 - Status quo on research data management
 - Desiderata
 - Enabling open science
 - Raise awareness
- Method:
 - Online survey (26 questions, German and English, anonymous)
 - Software *LimeSurvey*
 - Duration of survey: 19th January to end of March 2015
 - Analysis and report 4 months
 - Results published in German and English (November 2015)

A new ecosystem of services

High interest in support

60% Technical infrastructure

49% Specific support

42% Legal advice

41% Helpdesk

37% Training courses

Data types and formats

97% Text files

81% Graphics

67% Tables

34% Structured text

28% Video

27% Data base

23% Sourcecode

21% Audio

20% Software

8% Configuration files

Storage

Collaborative data infrastructures

Storage volume

per year

55% require
an average of
more than 50GB

7%
more than 1TB

11% No

de

software
generation files

Data Arch.

27%
after sign

Re-design of re-use scenarios

Are your research data

reusable
to others?

78% Yes or sometimes
22% No



Do you make
user agreements?

45% Yes:
Cooperation agreements
Open content licenses
Individual agreements

36% No

E
P
M

Towards an open science ecosystem

Incentives to **share** data

Increased visibility

New contacts/collaborations

Recognition

Relevance for evaluations



What's stopping you
from sharing?

Effort

Data abuse

Privacy violations

Misinterpretation, Falsification

Other legal restrictions

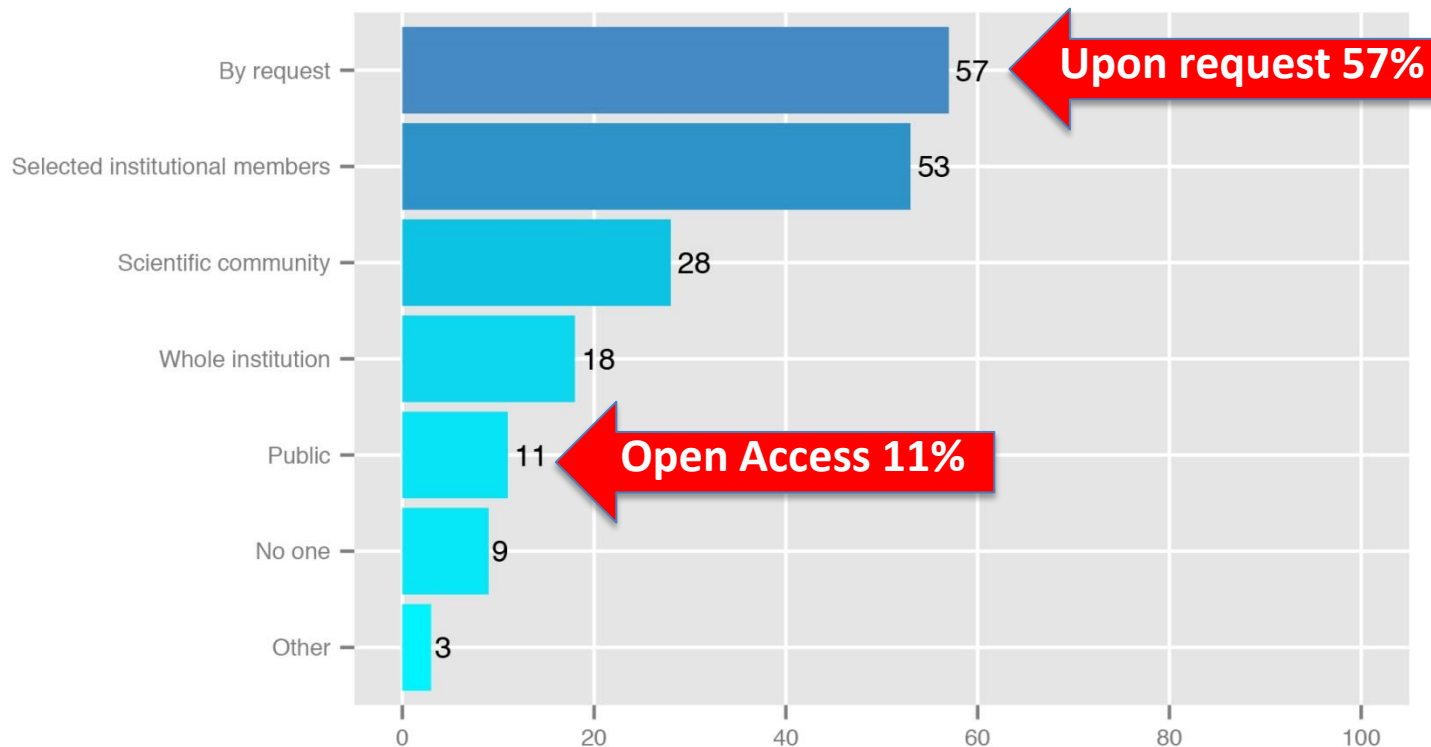
Commercial use

Increased competitive pressure



Identification of target groups

Whom do you grant access to your research data?



Relative response rate (%)

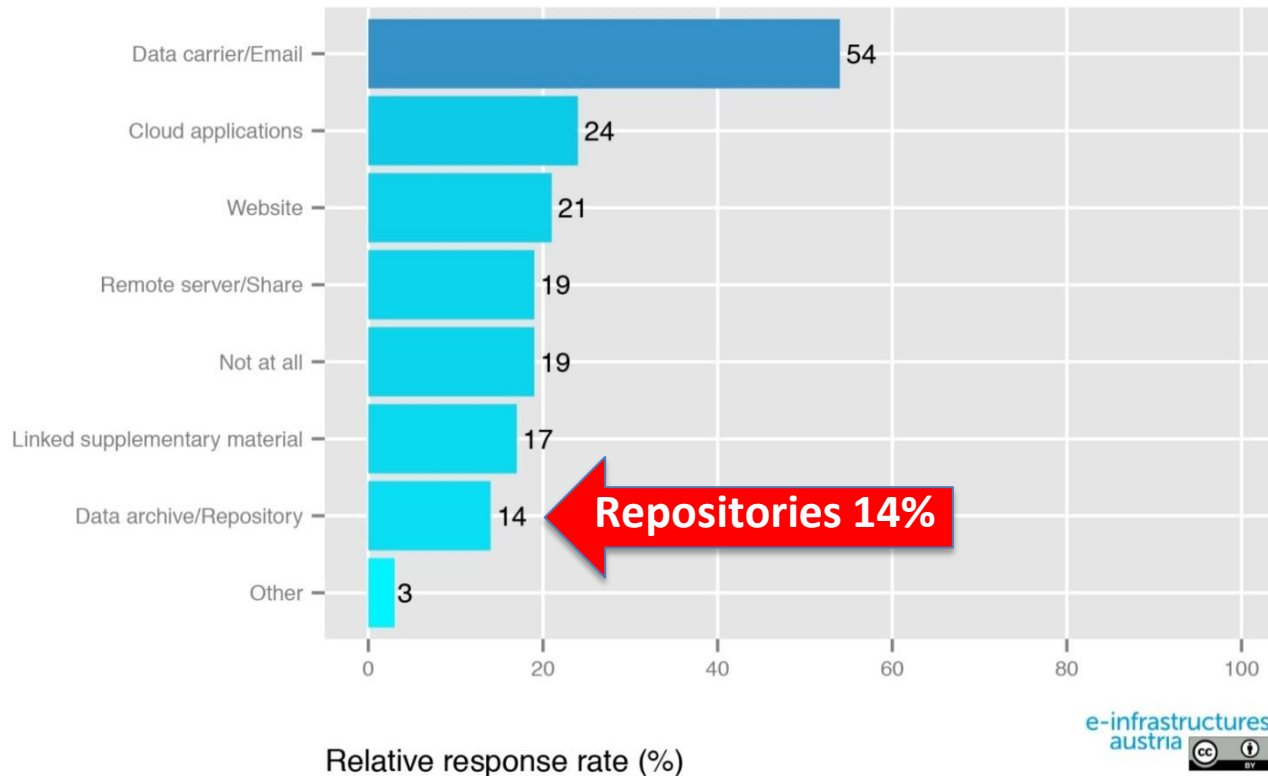
Re-thinking of scientific workflows

54% share their data by using
external storage devices or
email



Coordination between various e-infrastructure components

How can others access your research data?



Enhancing sustainability

ssing

What happens when you
leave your institution?

**43% Data remain
at institution**

36% Data are taken

5% Data are deleted

A common e-infrastructure umbrella



A common strategic vision

87% Configuration

Expectations
on the institution:
Qualified personnel
Guidelines
Policies

High interest in

Infrastr

References and downloads

www.e-infrastructures.at

Download full report:

Zenodo:

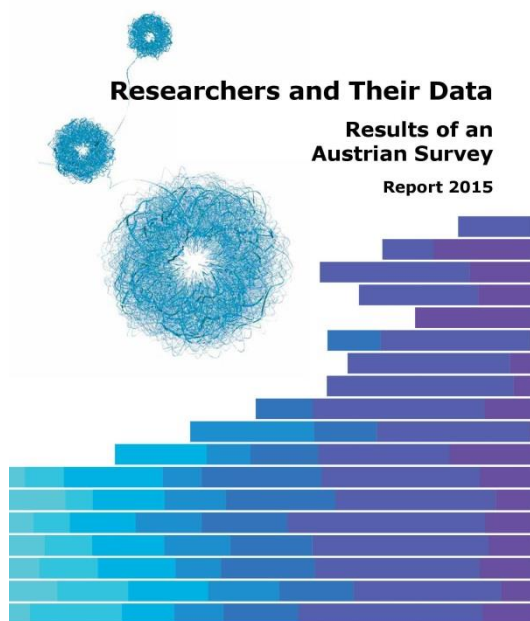
[DOI 10.5281/zenodo.34005](https://doi.org/10.5281/zenodo.34005)

Phaidra (e-book format):

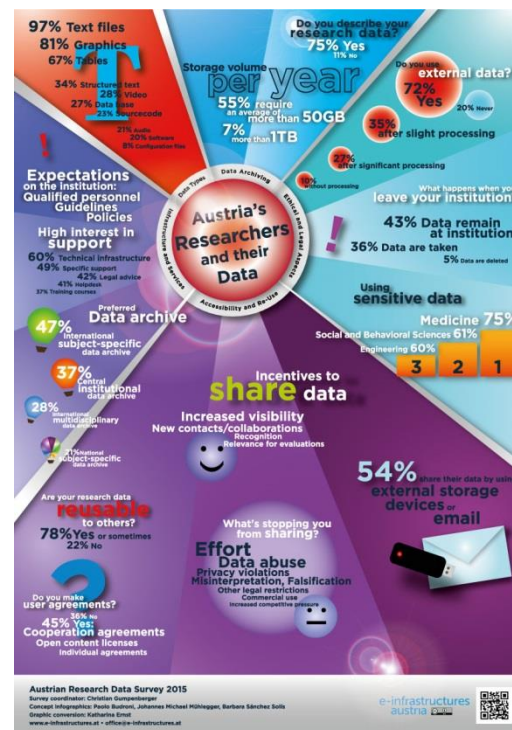
https://phaidra.univie.ac.at/detail_object/o:409473

Download Poster:

https://phaidra.univie.ac.at/detail_object/o:409452



e-infrastructures
austria



e-infrastructures
austria



FINDINGS

The findings from this survey form the basis for a consecutive

- a) **realization of RDM-policies**
- b) the identification of e-Infrastructure Commons
- c) the optimization of e-infrastructures and services available in this field, in accordance with needs that have been expressed
- d) the **realization of trainings** („essentials and data stewardship in e-infrastructures“).

About Open Science: The European Open Science Cloud

Open Science is the movement to
make scientific research, data and
dissemination accessible at all levels of
an enquiring society

Open Science

a paradigm shift in the modus operandi of research and science impacting the entire scientific process

Research Cycle

Analysis

Publication

Review

Conceptualization

Data Gathering

Characteristics

Citizen Science

Open code

Pre-print

Open Access

Alternative Reputation Systems

Collaborative Bibliographies

Science Blogs

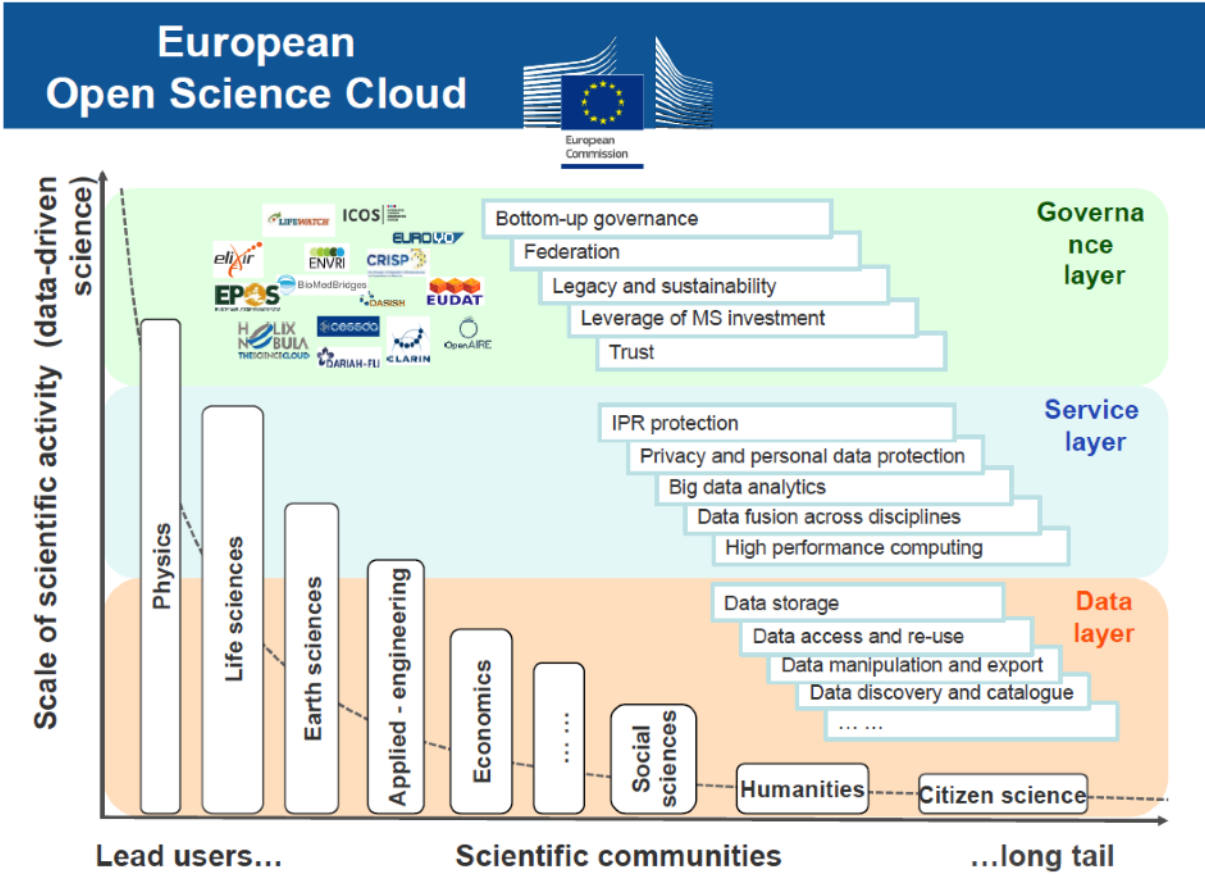
Open Annotation

Open Data

Open Lab Books/Workflows

Data Intensive

European Open Science Cloud



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

Realising the Open Science Cloud

On 11. Oktober 2016, the Commission High Level Expert Group on the EOSC (HLEG-EOSC) published its first report, entitled „Realising the Open Science Cloud.“ The report calls upon the **500,000 members** of the European research community to implement policies, construct effective governance models, identify e-Infrastructure commons, build up digital eco-systems, commit to data stewardship, **train data experts** and to define rules of engagement.

European Open Science Cloud

- Build on existing infrastructure and expertise
- Devise Rules of Engagement
- EU contribution to FAIR data and Open Science
- Build links to regional Cloud(s) in Latin America & Caribbean
- Develop expertise
 - Half a million ‘core data scientists’ in Europe
 - 5% of total research spend should be on data stewardship



Available [here](#)

³³EOSC - Key elements

- ★ Lightest possible, internationally effective governance
- ★ Guidance only where guidance is due: *greatest possible autonomy within scientific work clusters*
- ★ Rules of engagement for service provision: *introduce a governance for rules of engagement , projects and related teams*
- ★ Federate the gems and amplify good practice
- ★ Build on existing capacity and expertise where possible: *workshops, implementation of continuing-education seminars*
- ★ Optimize the e-Infrastructures communities: *constant feedback, competent servicing*

**Findings and conclusions.
About Policies → the Project
LEARN**

FINDINGS

Results are conform with international reviews

The findings from this survey form the basis for a consecutive

- a) realization of RDM-policies**
- b) the identification of e-Infrastructure Commons**
- c) the optimization of e-infrastructures and services available in this field, in accordance with needs that have been expressed**
- d) the realization of trainings („essentials and data stewardship in e-infrastructures“).**

Conclusions

- Results conform with international reviews
- Identification of challenges for enabling Open Science Vision
- Embedding in transnational e-infrastructures initiatives
- Deeper involvement of stakeholder groups – **activate them**
- **Reference points at local level**
- **Shared “vertical” services**
- **Horizontal services**
- **Need of know-how transfer of e-infrastructure essentials**
- **Release of RDM policies**

LEARN – LEaders Activating Research Networks

- Purpose is to develop the [LERU Roadmap for Research Data](#) to build a global co-ordinated global e-infrastructure
- Outputs
 - Model Research Data Management policy
 - Toolkit to support implementation
 - Executive Briefing in five core languages so as to ensure wide outreach

Started in June 2015; runs for 24 months
€497,000 budget - 100% funded

Horizon 2020

Call: H2020-INFRA supp-2014-2

Topic: INFRA supp-7-2014

Type of action: CSA

Proposal number: 654139

Proposal acronym: LEARN

LEARN: target audience

→ deeper involvement of stakeholder group

- **Researchers**

↳ LEARN will support effective data management planning

- **Research institutions**

↳ LEARN will provide a model RDM policy

- **Data scientists/curators**

↳ LEARN will help to identify systems, infrastructure and policy

- **Liaison and Support services**

↳ LEARN will help staff understand researcher RDM needs

- **Research funders**

↳ LEARN will help to inform funder policy on RDM

The LERU Roadmap for Research Data

- A guide for **universities and research organisations** on how to engage with the potential and tackle the challenges of data-driven research
- Published **December 2013**
- Available at: <http://goo.gl/INWlcl>



LEARN will:

Communicate the LERU Roadmap for Research Data to an international community

Raise awareness in Research Data Management (RDM) issues and policy

Stimulate further research on RDM issues

Encourage the production of RDM policies at an institutional level

Thank you

Dr Paolo Budroni

University of Vienna

LEARN - Policy Development and Alignment

e-Infrastructures Austria | www.e-infrastructures.at

Open Education Austria | www.openeducation.at

e-Infrastructure Reflection Group,

Austrian National Delegate | www.e-irg.eu

e-mail: paolo.budroni@univie.ac.at