Open Science – Open Data – Open Services

Open Science Commons/Cloud?

Per Öster

CSC – IT Center for Science Ltd

www.eudat.eu

EUDAT receives funding from the European Union's Horizon 2020 programme - DG CONNECT e-Infrastructures. Contract No. 654065

Science as an open enterprise

June 2012

THE ROYAL SOCIETY

SUMMARY

Recommendations

This report analyses the impact of new and emerging technologies that are transforming the conduct and communication of research. The recommendations are designed to improve the conduct of science, respond to changing public expectations and political culture and enable researchers to maximise the impact of their research. They are designed to ensure that reproducibility and self-correction are maintained in an era of massive data volumes. They aim to stimulate the communication and collaboration where these are needed to maximise the value of data-intensive approaches to science. Action is needed to maximise the exploitation of science in business and in public policy. But not all data are of equal interest and importance. Some are rightly confidential for commercial, privacy, safety or security reasons. There are both opportunities and financial costs in the full presentation of data and metadata. The recommendations set out key principles. The main text explores how to judge their application and where accountability should lie

Recommendation 1

Scientists should communicate the data they collect and the models they create, to allow free and open access, and in ways that are intelligible, assessable and usable for other specialists in the same or linked fields wherever they are in the world. Where data justify it, scientists should make them available in an appropriate data repository. Where possible, communication with a wider public audience should be made a priority, and particularly so in areas where openness is in the public interest.

Although the first and most important recommendation is addressed directly to the scientific community itself, major barriers to widespread adoption of the principles of open data lie in the systems of reward, esteem and promotion in universities and institutes. It is crucial that the generation of important datasets, their curation and open and effective communication is recognised, cited and rewarded. Existing incentives do not support the promotion of these activities by universities and research institutes, or by individual scientists. This report argues that universities and research institutes should press for the financial incentives that will facilitate not only the best research, but the best communication of data. They must recognise and reward their employees and reconfigure their infrastructure for a changing world of science.

Here the report makes recommendations to the organisations that have the power to incentivise and support open data policies and promote data-intensive science and its applications. These organisations increasingly set policies for access to data produced by the research they have funded. Others with an important role include the learned societies, the academies and professional bodies that represent and promote the values and priorities of disciplines. Scientific journals will continue to be media through which a great deal of scientific research finds its way into the public domain, and they too must adapt to and support policies that promote open data wherever appropriate.

Recommendation 2

Universities and research institutes should play a major role in supporting an open data culture by: recognising data communication by their researchers as an important criterion for career progression and reward; developing a data strategy and their own capacity to curate their own knowledge resources and support the data needs of researchers; having open data as a default position, and only withholding access when it is optimal for realising a return on public investment.

Recommendation 3

Assessment of university research should

SUMMARY

Recommendation 5

Research Councils and Charities should improve the communication of research data from the projects they fund by recognising those who could maximise usability and good communication of their data; by including the costs of preparing data and metadata for curation as part of the costs of the research process; and by working with others to ensure the sustainability of datasets.

Recommendation 6

As a condition of publication, scientific journals should enforce a requirement that the data on which the argument of the article depends should be accessible, assessable, usable and traceable through information in the article. This should be in line with the practical limits for that field of research. The article should indicate when and under what conditions the data will be available for others to access.

Effective exchange of ideas, expertise and people between the public and private sectors is key to delivering value from research. The economic benefit and public interest in research should influence how and when data, information and knowledge from publicly or privately funded research are made widely available.

Recommendation 7

Industry sectors and relevant regulators should work together to determine the approaches to sharing data, information and knowledge that are in the public interest. This should include negative or null results. Any release of data should be clearly signposted and effectively communicated.

Recommendation 8

Governments should recognise the potential of open data and open science to enhance the excellence of the science base. They should develop policies for opening up scientific data that complement policies for open government data, and support development of the software tools and skilled personnel that are vital to the success of both.

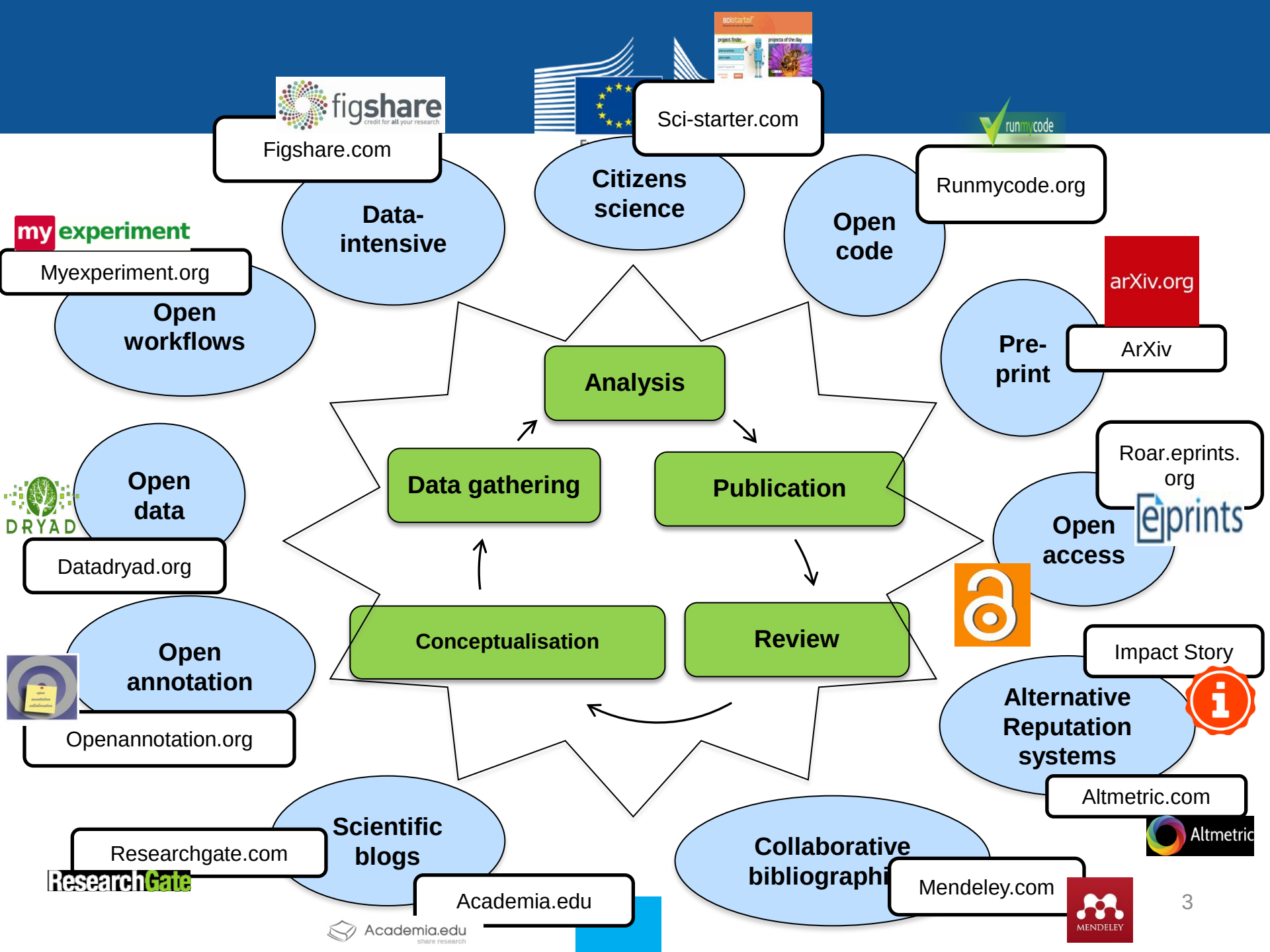
Judging whether data should be made more widely available requires assessment of the public benefits from sharing research data and the need to protect individual privacy and other risks. Guidance for researchers should be clear and consistent.

Recommendation 9

Datasets should be managed according to a system of proportionate governance. This means that personal data is only shared if it is necessary for research with the potential for high public value. The type and volume of information shared should be proportionate to the particular needs of a research project, drawing on consent, authorisation and safe havens as appropriate. The decision to share data should take into account the evolving technological risks and developments in techniques designed to safeguard privacy.

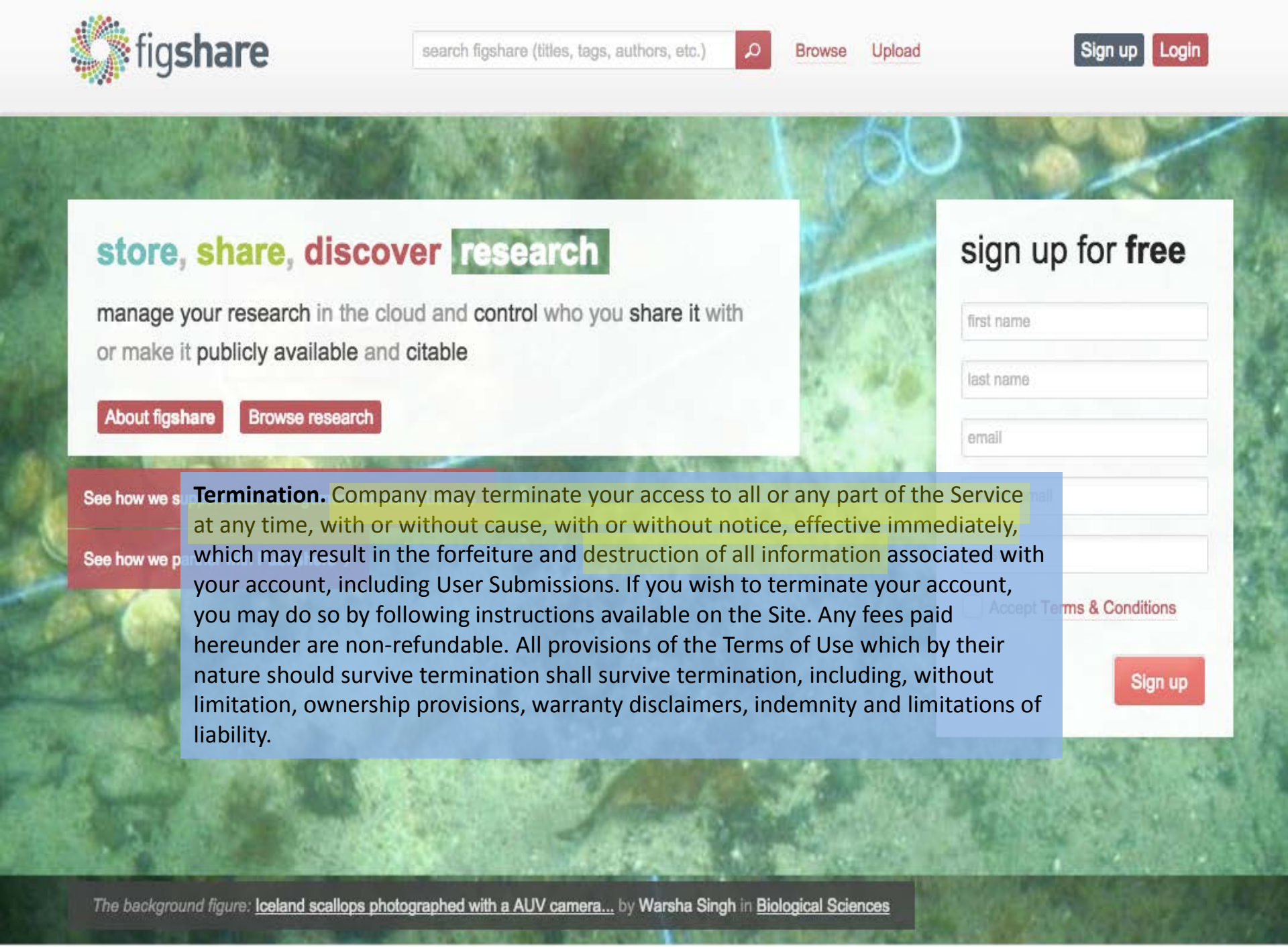
Recommendation 10

In relation to security and safety, good practice and common information sharing protocols based on existing commercial standards must be adopted more widely. Guidelines should reflect the fact that security can come from greater openness as well as from secrecy.


Figshare.com
Sci-starter.com
Runmycode.org
Data-intensive
Citizens science
Open code
Myexperiment.org
Open workflows
arXiv.org
ArXiv
Pre-print
Analysis
Data gathering
Publication
Open data
Datadryad.org
Roar.eprints.org
Open access
Conceptualisation
Review
Open annotation
Openannotation.org
Impact Story
Alternative Reputation systems
Altmetric.com
Scientific blogs
Researchgate.com
Collaborative bibliographi
Mendeley.com
Academia.edu

Integration?

Sustainability?



Termination. Company may terminate your access to all or any part of the Service at any time, with or without cause, with or without notice, effective immediately, which may result in the forfeiture and destruction of all information associated with your account, including User Submissions. If you wish to terminate your account, you may do so by following instructions available on the Site. Any fees paid hereunder are non-refundable. All provisions of the Terms of Use which by their nature should survive termination shall survive termination, including, without limitation, ownership provisions, warranty disclaimers, indemnity and limitations of liability.

Integration

- **Integration** (EUDAT CDI)
  - compose & combine EUDAT technical services
    - ➜ “B2 Enterprise Edition”
- **Integration** (European e-infrastructure)
  - agree protocols, interfaces, identity management with HPC, cloud, and networks
    - ➜ open EU data/compute platform for research
- **Integration** (European Research)
  - agree policies, API, and methods with universities, libraries, digital publication actors, and service companies
    - ➜ open science

Sustainability

- **Sustainable** (EUDAT CDI)
  - create partnership of sustainable organizations and develop it
    - ➜ utilise the full potential of “B2 Enterprise Edition”
- **Sustainable** (Financial)
  - ensure multiple revenue streams for partners and partnership
    - ➜ open EU data/compute platform for research
- **Sustainable** (Societal)
  - follow policies of governments, universities, libraries, digital publication authorities
    - ➜ open science, open society

Vision

30th October 2015

Position Paper: European Open Science Cloud for Research

Summary

As part of the Digital Single Market strategy,[1] the *Open Science Cloud* will raise research to the next level. It promotes not only scientific excellence and data reuse but also job growth and increased competitiveness in Europe, and drives Europe-wide cost efficiencies in scientific infrastructure through the promotion of interoperability on an unprecedented scale. The Open Science Cloud offers researchers from all disciplines seamless, open access to the advanced digital capabilities, resources and expertise they need to collaborate and to carry out data- and computing-intensive science. Secure and trustworthy, the Open Science Cloud engages researchers in governing, managing and preserving resources for everyone's benefit. The Open Science Cloud is an open, service-driven endeavour, inclusive of all stakeholders. Governed as a commons, it leverages two decades of public and private investment in e-infrastructures for the benefit of scientific research and innovation.

Background

Science is changing, both in the way it is performed and the way it is communicated. Driven by remarkable advances in information and communication technologies, today's scientific infrastructures offer researchers unprecedented access to data sources, data-intensive sensors, and increasingly comprehensive analysis and simulation facilities that have revolutionized scientific methods in a remarkably short space of time. Research services, processes and outputs are becoming accessible to all levels of society. Enormous amounts of data are being generated, bringing extraordinary new opportunities for their innovative reuse in novel scientific, commercial, and citizen-science contexts. This is *Open Science*.

Open Science is a key driver, not only of scientific progress, but also of economic and societal innovation. To harness its full value and reap the fruits of public and private investment, Europe needs to foster an open, collaborative platform for the management, analysis, sharing, reuse and preservation of research data on which innovative services can be developed and delivered. For this, Europe can build on decades of public investment in scientific infrastructures—experimental facilities, networking, high-performance and high-throughput computing, cloud services, scientific software and institutional and community data repositories—by connecting national and international infrastructures and services. The Open Science Cloud is the vehicle to achieve this vision. Below we articulate the eight essential elements it needs to succeed.

Many of the resources and services needed for the Open Science Cloud already exist; while technical challenges remain, most of the barriers are ones of policy and concern funding, lack of interoperability, access policies and coordinated provisioning. The Open Science Cloud will address these issues and enrich and further advance the portfolio of resources and services to make the entire scientific lifecycle more open and transparent. To this end, governance of the Open Science Cloud will be modelled after the governance of the Internet, conducted by a decentralized, international group of stakeholders drawn from across research and civic society, from both public and private sectors. The Open Science Cloud's governance will hold custody of the shared services, policies and standards that maintain its persistency, its global interoperability and its adherence to the Open Science vision. By involving all the relevant stakeholders who support today's research—funding agencies, policy makers, research infrastructures, e-Infrastructures, libraries, data providers and service providers—the Open Science Cloud will significantly impact the way research is done in Europe and will put European research at the forefront of Open Science globally.

[1] http://ec.europa.eu/digital-agenda/en/news/open-science-competitiveness-council-28-29-may-2015

1

30th October 2015

The Open Science Cloud: Eight Elements for Success

I. **Open:** This is the driving principle of the Open Science Cloud: openness in design, in participation and in use. The Open Science Cloud will be based on open access and promote the development and adoption of open standards, enabling collaborative environments with no artificial barriers to participation or resource-sharing by any stakeholder. It will enable accessibility, transparency, and reproducibility in all stages of the research life-cycle. Having a flexible open design, the Open Science Cloud will foster public-private partnerships, turning all investment into economic growth.

II. **Publicly funded & governed:** A publicly funded and publicly governed Open Science Cloud will guarantee persistence and sustainability, and ensure that outcomes are driven by scientific excellence and societal needs rather than profit. This "commons" approach, welcoming partnership with private-sector actors while driven by the public good, will encourage the development of innovative services that are conducive to the future of Open Science, while guaranteeing the long-term, persistent care of resources.

III. **Research-centric:** Following the true spirit of agile co-design and participation, researchers and research communities—including those from the private sector—will be fully engaged in the design of the Open Science Cloud, to ensure the development of services responsive to their needs.

IV. **Comprehensive:** The Open Science Cloud will be universal, specific to no single scientific discipline or research field. It will promote inter- and multi-disciplinary science and encourage innovation and integrated knowledge creation among all research communities, also capturing the long tail of science and citizen science.

V. **Diverse & distributed:** The Open Science Cloud will leverage the richness of Europe's distributed e-infrastructures, encompassing a resilient network of actors, resources and services organized nationally and at the European level. Embracing diversity through openness, the Open Science Cloud will drive a more efficient use of ICT investments across infrastructures and communities, addressing the digital divide and lowering the barriers to adoption for institutions and researchers.

VI. **Interoperable:** Through the promotion and adoption of common standards and protocols for all resources and digital services, the Open Science Cloud will connect networks, data, computing systems, software, tools and services for research as seamlessly as the Web connects information.

VII. **Service-oriented:** The Open Science Cloud will be protocol-centric and service-oriented. It will provide services that address the full research lifecycle, including data gathering, management, analysis, sharing and discovery. The Open Science Cloud will be the framework and testing environment for new, innovative methodologies and services that further advance research in the Open Science context.

VIII. **Social:** The Open Science Cloud will be a socio-technical endeavour that connects diverse communities and promotes the development of human networks. By adopting community-based rules and procedures with incentives for sharing and responsible use, it will enable the sharing of knowledge and facilitate the embedding of Open Science practices into researchers' everyday workflows. This will require a strong social dimension of consultation, outreach, advocacy, training and support, in an ecosystem of local, national and international programmes.

Dr. Kimmo Koski, Project Coordinator, EUDAT
Kristiina Hormia-Poutanen, President, LIBER
Prof. Mike Chatzopoulos, Project Coordinator, OpenAIRE
Yannick Legré, Director, EGI
Dr. Bob Day MBE, Chief Executive, GÉANT

2

http://dx.doi.org/10.5281/zenodo.32915