Title: Share, steward and reuse research data - European Open Science Cloud

Author(s): Södergård, Caj

Citation: 29th NORDUnet Conference, 20 - 22 September 2016, Helsinki, Finland.
NORDUnet A/S (2016)

Date: 2016

Rights: This presentation may be downloaded for personal use only.

By using VTT Digital Open Access Repository you are bound by the following Terms & Conditions.

I have read and I understand the following statement:

This document is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of this document is not permitted, except duplication for research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered for sale.
Share, steward and reuse research data
- European Open Science Cloud

NORDUnet 2016

Prof. Caj Södergård
Content

- Why EOSC ?
- What is EOSC ?
- What does the High Level Expert Group propose ?
- What next ?
Towards Data-driven Research

- Big data needs cloud, bandwidth and powerful computers -> e-infrastructures have to adapt

- One paper is published every 30 seconds (2014).
  - 70,000 papers have been published on a single protein, the tumour suppressor p53

- A considerable share (10 - 90 %) of scientific papers are never cited, nor even read.
  - Pre-clinical oncology – 89% not reproducible (Nature 2012)

- A growing pressure to share scientific data
Science 2.0

"How can I use participative Internet technologies in my research?"

- Prepare project applications collaboratively
- Blog about my research
- Discuss it on social networks (e.g. Twitter)
- Exchange data online

Must NOT be open, but makes the transition to Open Science easier

Adapted from Guido Scherp, 2016
Photo: Flickr
**Science 2.0 and Open Science**

**Science 2.0**

"How can I use participative Internet technologies in my research?"

- Prepare project applications collaboratively
- Blog about my research
- Discuss it on social networks (e.g. Twitter)
- Exchange data online

*Must NOT be open, but makes the transition to Open Science easier*

**Open Science**

"How can I make my research as public and transparent as possible?"

- Publish in Open Access
- Publish data (including raw data)
- Publish scripts, source code
- Licence for reuse (e.g. CC-0)
- Have transparent processes (reviews etc)
- Tell publicly about the research and future ideas

*Works without Science 2.0 – but Science 2.0 gives support*

"Knowledge is open if anyone is free to access, use, modify, and share it …" [Open Definition]

Adapted from Guido Scherp, 2016

Photo: Flickr
Cloud Initiative
Science in transition - digitalization/data-centric - infrastructures, practices and methods must also change.

The European Cloud Initiative will make it easier for researchers and innovators to find, access, interoperate and re-use data, and will reduce the cost of data storage and high-performance analysis.

Making research data openly available can help boost Europe’s competitiveness by benefitting start-ups, SMEs and data-driven innovation, including in the fields of medicine and public health. It can even spur new industries, as demonstrated by the Human Genome Project.

**OPPORTUNITIES**

- Spurring new solutions in complex areas like #eHealth, transport, environment
- Better public services such as #smartcities
- Better science for complex problems
- Commercial opportunities for innovative companies

Better value for taxpayers – opening up data produced by projects funded by the Horizon 2020 research and innovation programme: Findable, Accessible, Interoperable, Reusable
EU and industry will invest 6,7 B€

- **2016:** creating a European Open Science Cloud (EOSC)
  - First calls
    - INFRADEV-04-2016 call
    - EINFRA-12-2017 call

- **2017:** opening up by default all scientific data produced in the €77 billion Horizon 2020

- **2018:** launching the quantum technology flagship

- **2020:** developing and deploying a large scale European high performance computing, data storage and network (EDI)

Both the EOSC and the EDI will build on existing EU-funded e-infrastructures

WHAT is European Open Science CLOUD?
Our goal is to create a European Open Science Cloud to make science more efficient and productive and let millions of researchers share and analyse research data in a trusted environment across technologies, disciplines and borders…"
- European
- Open
- Science
- Cloud
What does the High Level Expert Group propose?
High Level Expert Group on EOSC

• The Commission will publish the HLEG report in September 2016
  • Contains initial guidelines for how to go forward towards a European Open Science Cloud

• HLEG Members
  • Barend Mons (Chair)
  • Paul Ayris
  • Jean-Yves Berthou
  • Rachel Bruce (Rapporteur)
  • Stefanie Lindstaedt
  • Anna Monreale
  • Yasuhiro Murayama (Observer, Japan)
  • Caj Södergård
  • Klaus Tochtermann
  • Ross Wilkinson (Observer, Australia)
Draft report: Key EOSC requirements

- New modes of scholarly communication
- Modern reward and recognition practices
- Core data scientists need to be trained and their careers supported
- Cross-disciplinary collaboration
Report: Key EOSC requirements

- Fostering transition from Science to Innovation
- A complex eco-system of infrastructures
- Machine understanding
- Findable, Accessible, Interoperable, Reusable (FAIR)
EOSC: Policy Recommendations

- Take actions with Member States
- Build on existing capacity and expertise
- Support Global Research Data Commons, with open protocols
EOSC: Governance Recommendations

- Aim at light, internationally effective governance
- Guidance only where guidance is due
- Define Rules of Engagement for participation in EOSC
EOSC: Implementation Recommendations

- Develop, Endorse and Implement Rules of Engagement

- Develop a plan plan for governance of EOSC

- Fund an effort to locate and develop Data Expertise in Europe

- Install innovative funding scheme for preparatory phase

- Make adequate Data Stewardship mandatory for all research proposals
  - 5% of budget has been discussed

- Install executive teams for preparation and international coherence
How forward – steps discussed

1. A **Memorandum of Understanding** (MoU) between the major players in the Member States and the international coordination level

2. **Rules of Engagement** for public and private parties who want to play an active role in the implementation of the EOSC ‘national nodes’

3. A further detailing of how the ‘**Cloud Coin**’ (dedicated Data Stewardship funding to be spent with ‘certified’ or ‘conformant’ EOSC providers’) could be implemented.

4. **Skill determination** for core data experts/data stewards and the building of a coordinated curriculum and training materials (possibly together with WG on Open Education)

5. **Support tools** to assist researchers in the development of appropriate Data Management Plans (possibly together with FAIR Data WG)
The Cloud Coin Model

1. Approves DS Plan before submission of grant
2. Requests Credits
3. Selection
4. Delivers Funding Recommendation
5. Review & Approval
6. Directs reseller to distribute credits
7. Distributes
8. Uses credits

Endorsed Repository
Data Publisher
Cloud Provider A
Cloud Provider B
Cloud Provider C
HPA service

Investigator Institution

In DS plan

Monthly invoice to Cloud Coin Authority?

Cloud Coin Authority

FunDer

Review panel (DS = Tickbox)
Thank you for your attention