Policies and principles of Open Scholarship: EC perspective

Pre-RDA plenary Workshop on Data and computing infrastructures for open scholarship

Paris, 22 September 2015

Jarkko Siren
European Commission
DG CONNECT
eInfrastructure

Author’s views do not commit the European Commission
Open Scholarship and the EC

**EC policies**

- Open Access (publications and data) and Open Science
- 'Open Scholarship' implied (opportunity?)

**Definitions**

- Research methods, scholarly communication, scholarly activity in general

**Focus on infrastructures as enablers of Open Scholarship**
Analysis

Data-gathering

Publication

Conceptualisation

Review

Citizens science

Open code

Pre-print

Open access

Alternative Reputation systems

Open workflows

Open data

Open annotation

Scientific blogs

Collaborative bibliographies

Open code

Data-intensive

Towards better access to scientific information: boosting the benefits of public investments in research, COM(2012) 401 final - July 2012

Commission, Recommendation on access and preservation of scientific information, C(2012) 4890 final – July 2012

Horizon 2020
- Open Access to Scientific Publications
- Pilot on research data: Data Management Plan
Open Science
Competitiveness Council 29 May 2015

Member States have expressed their wish for the development of a European Open Science Agenda

Council Conclusion, 29 May 2015:

CALLS for action to remove obstacles to wide access to publicly funded research publications and underlying data;
CALLS for actions addressing better data management and, in this context, WELCOMES the Pilot on Open Research Data under Horizon 2020;

In the context of the implementation of the European Research Area (ERA), LOOKS FORWARD to the possible development of action plans or strategies for open science
OA to publications mandate in H2020

Each beneficiary must ensure OA to all peer-reviewed scientific publications relating to its results:

- **Deposit** published version or final peer-reviewed manuscript in a repository of the researchers choice
- Ensure OA on publication or at the latest within 6/12 months
- Ensure OA to the bibliographic metadata that identify the deposited publication, via the repository
- **Aim to deposit** at the same time the research data needed to validate the results ("underlying data")
Pilot on Open Research Data in H2020

It is pilot: a way to define future action.

Types of data concerned:

• Data needed to validate the results presented in scientific publications ("underlying data")

• Other data as specified in Data Management Plan (=up to projects)

• Data Management Plans (DMPs) mandatory for all projects participating in the pilot
Pilot on Open Research Data in H2020

Beneficiaries participating in the Pilot will:

- Deposit this data in a research data repository of their choice
- Take measures to make it possible to access, mine, exploit, reproduce and disseminate free of charge
- Provide information about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (where possible, provide the tools and instruments themselves)
Horizon 2020

- Over €70 billion research and innovation funding programme (2014-2020)
- Includes Capacity building for Open Scholarship through e-Infrastructures: €100 in 2014-15 and €200 in 2016-17
- e-infrastructures WP 2016-17 (draft version online):
  - Data and computing e-infrastructures for Open Science
  - Repository networks, publishing platforms, digital identifier infrastructures, open review, certification, etc.
Staying Competitive in Science

- Large scale collaborations becoming the norm
  - *often global*
  - *virtual research communities*
  - *access to rare/remote resources*

- Data-intensive science and innovation
  - *Use and manage exponentially growing sets of data*

- Experimentation in silico, simulation
  - *Use of high-performance computing*
3 guiding principles

- Data and Computing e-infrastructures go together
  - Research Infrastructures and e-Infrastructures go together
- Research Data – European Policy Framework
- Research Data Alliance
  - Report “Data Harvest” (follow-up of “Riding the Wave”)
Research logic machines

Now research data is stored in digital form. Easier to be processed by "logic machines" programmed with complex models able to dig into the data.

Logic machines are made of human scientific knowledge and creativity, software and the underlying hardware.

Scientist notebooks can now be linked to a huge amount of other data resources (including scientific papers), computers with unprecedented capacity, eventually connected to global networks.
Digital scholarly record

Repositories for publications, data, software, etc. could play an increasingly important role.

+ identifier infrastructures + registries

Open Scholarship requires new designs and architectures. Some of them might exist already.
Thank you!

Jarkko Siren
jarkko.siren(at) ec.europa.eu