Open Research Data and Open Science: The European Perspective

Celina Ramjoué

Head of Sector, OA to scientific publications and data

European Commission

DG Communications Networks, Content and Technology (CONNECT) - Digital Science Unit

Opendata.ch 2016

Lausanne, 14 June 2016

Outline

- Framing
- Open Research Data Pilot
- Experiences
- Next steps

Outline

- **Framing**
- Open Research Data Pilot
- Experiences
- Next steps

Framing I: Digital Single Market (DSM)

DSM: "Market in which free movement of goods, persons, services and capital is ensured and where individuals and businesses can seamlessly access and exercise online activities."

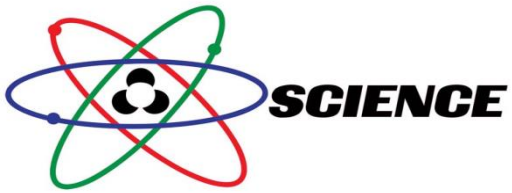
One of ten **Juncker priorities**

One of the DSM pillars: focus on maximising growth potential of the digital economy by building a **data economy**

Themes: Copyright, including text and datamining (TDM), open science, free flow of data, open science cloud

Framing II: Open science

- **Open science is … the transformation and opening up of science, research and innovation through Information and Communication Technologies, with the objective of making science more efficient, transparent and interdisciplinary, and of enabling broader societal impact and innovation.**
- **Areas: open access, citizen engagement, open e-infrastructure, research metrics, etc.**
- **Commissioner Moedas: open science, open innovation, open to the world**



Member States: Council Conclusions on the data-driven economy (29 May 2015)

Member States emphasise the data-driven economy and support open science

The Council:

- RECOGNISES the high potential of the **data-driven economy**.
- REAFFIRMS the broad political support from Member States for setting better framework conditions for faster and wider **data-driven innovation taking into account the research perspective.**
- LOOKS FORWARD to the possible development of action plans or strategies for **open science**.

Member States: Council Conclusions on open science (27 May 2016)

Member States support open science

The Council:

- ACKNOWLEDGES that open science has the potential to **increase the quality, impact and benefits of science** and
- to accelerate advancement of knowledge by making it […] **better understandable by society and responsive to societal challenges**, and
- has the potential to **enable growth and innovation** through reuse of scientific results by all stakeholders at all levels of society, and
- ultimately **contribute to growth and competitiveness** of Europe.

Outline

- Framing
- **Open Research Data Pilot**
- Experiences
- Next steps

Open Research Data Pilot in H2020

- Development
- What data?
- Requirements
- Opting out
- Data management and DMPs

Development

- Launch: December 2013 with Horizon 2020
- Basis: open access to publications (mandate in Horizon 2020)
  - Requirement: aim to deposit underlying data
  - Text is data
- 2014-15 H2020 Work Programme: 7 areas (with opt-in possibility)
- 2016-2017 H2020 Work Programme: 9 areas (with opt-in possibility)
- As of 2017: all H2020 areas in the ORD Pilot, opt-out only (2016 announcement)

Pilot on ORD: what data?

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)

Pilot on ORD: requirements

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)

Approach: as open as possible, as closed as necessary

Source: Model Grant Agreement art 29.3

Pilot on Open Research Data: Opting out

Projects may opt out of the Pilot on Open Research Data in Horizon 2020 in a series of cases (submission stage):

- If the project will not generate / collect any data
- Conflict with obligation to protect results
- Conflict with confidentiality obligations
- Conflict with security obligations
- Conflict with rules on protection of personal data
- If the achievement of the action's main objective would be jeopardised by making specific parts of the research data openly accessible (to be explained in data management plan)

Opting out during project also possible

Being in the Pilot does not mean opening all data

Data management in Horizon 2020

- Brief **data management section** for all project submissions → part of evaluation
- **Data Management Plans (DMPs)** mandatory projects participating in the Pilot, optional for others

  → not part of evaluation
- DMPs to be generated within first 6 months of project, updates as needed (typically at time of review and at project end)

Data Management Plan (DMP)
(H2020 Data Management Guidelines, annex 1)

- For each dataset
- What data will be collected / generated?
- What standards will be used / how will metadata be generated?
- What data will be shared/opened and/or exploited?
- How will data be curated and preserved?

DMP: tool to determine what datasets can/cannot be open

Resources

H2020 Guidelines (2015 versions):

Open access (publications, data):
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Data management:
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Resources

OpenAire https://www.openaire.eu/

- Based on FP6, FP7, H2020 projects
- Implements FP7 and H2020 open access policies ...
- and much more ...
- Federated network of Open Access suject-based and institutional repositories and archives
- Repository: http://zenodo.org/
- OpenAIRE gives access to some 13 million publications and some 16 thousand datasets
- Network of National Open Access Desks (NOADs)

Outline

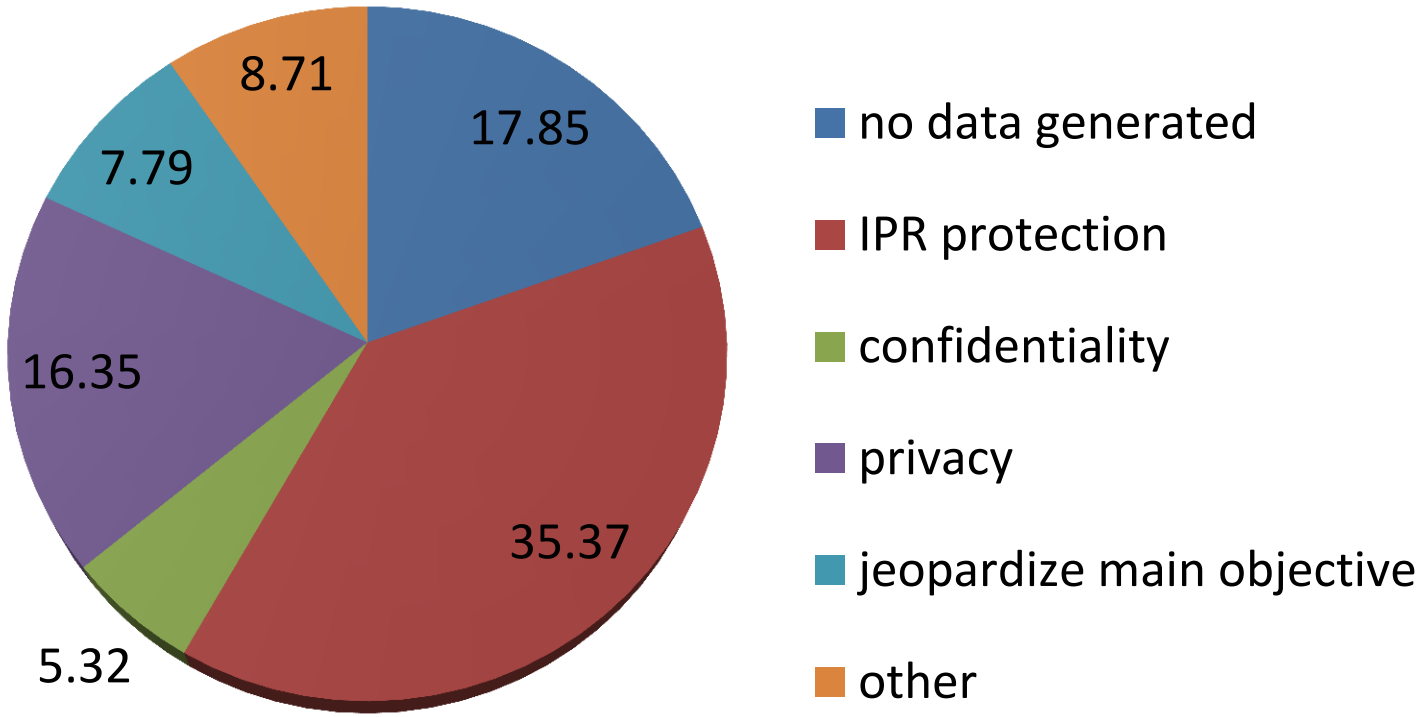
- Framing
- Open Research Data Pilot
- **Experiences**
- Next steps

ORD Pilot: take-up in first calls of H2020

- Basis: 3699 Horizon 2020 signed grant agreements (summer 2015)
- Calls in core-areas: **opt out 34,6%**

  → 65,4% of projects in the core areas participate in the Pilot
- Other areas: voluntary **opt in 11,9%**

ORD Pilot: opt-out reasons among proposals



The ORD Pilot data is now online (Open Data Portal)

Data: http://data.europa.eu/euodp/repository/ec/dg-rtd/ORD/Open_Access_Research_Data_Pilot.xlsx

Explanations:
https://data.europa.eu/euodp/data/dataset/open-research-data-the-uptake-of-the-pilot-in-the-first-calls-of-horizon-2020

2015 data, new dataset planned for autumn 2016

ORD Pilot: experiences (1)

- Explanation is paramount!
- Misperception that 'open' bias is evaluated positively.
- Confusion: DMP versus data management section at submission stage.
- Emphasise flexibility (many opt-out/opt-in mechanisms).
- It helps to re-frame ORD Pilot as Data Management Pilot.
- Need to state that not everything must be open. In theory, it is possible to be in the ORD Pilot and not open any data!
- At the same time: too easy to opt out (e.g. privacy)? (→ privacy / data protection by design approach needed?)

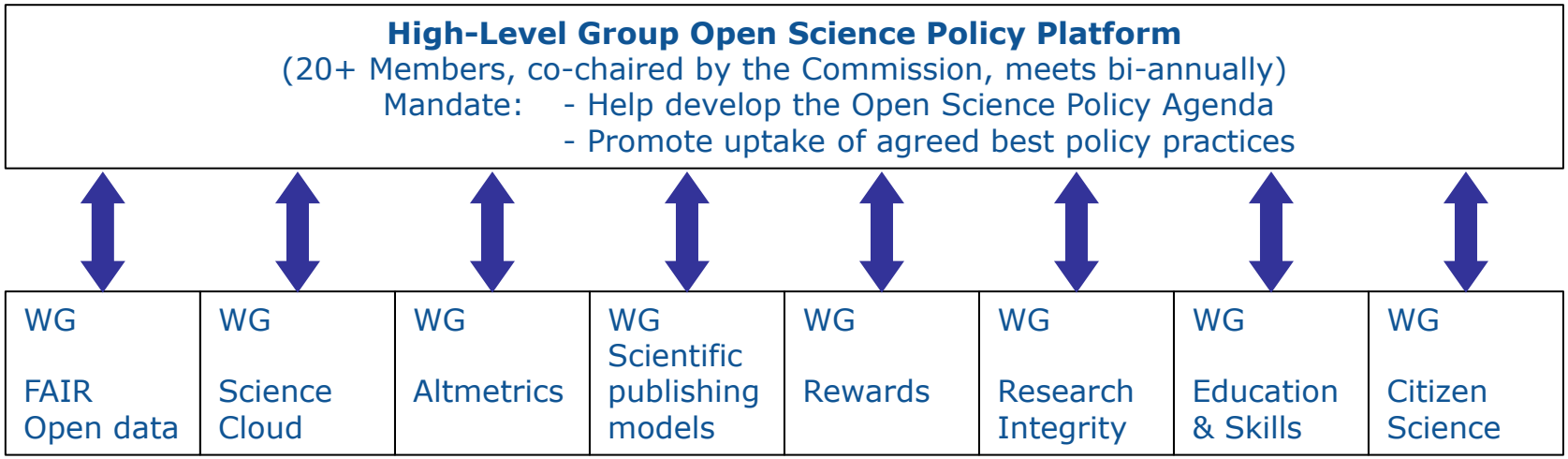
ORD Pilot: experiences (2)

- Stress the fact that researcher has freedom and responsility via DMP. Excellent research must include excellent data management.
- Questions about eligibility of data management costs.
- Tools and support needed for data management / DMPs.
- Emphasise the importance of feedback for policy the next Framework Programme: being in the Pilot means co-shaping European policy on opening up research data (Midterm review).
- Underline overall aim: kick-starting a virtous circle and change of culture.

Outline

- Framing
- Open Research Data Pilot
- Experiences
- **Next steps**

Open Science Policy Platform



High Level Group members:

http://ec.europa.eu/research/openscience/pdf/ospp_nominated_members.pdf#view=fit&pagemode=none

Announced 27 May 2016

European Cloud Initiative

Source: COM(2016)178 – 19 April 2016

Content: European Open Science Cloud, European Digital Infrastructure, Widening the user base (e-government & industry) → **about ALL data!**

European Open Science Cloud

- A virtual environment for all European researchers to store, manage, analyse and re-use data
- Bringing together existing and emerging data infrastructures
- Added value: scale, data-driven science, inter-disciplinarity, data to knowledge to innovation

Basis: builds on long-time funding and policy work in e-Infrastructure and cloud computing

Specific new action relevant to the ORD Pilot: The Commission will make open research data the default option, while ensuring opt-outs, for all new projects of the Horizon 2020 programme.

Thank you!

Celina Ramjoué

Head of Sector, OA to scientific publications and data

European Commission

DG Communications Networks, Content and Technology (CONNECT) - Digital Science Unit

Opendata.ch 2016

Lausanne, 14 June 2016