Opening up Scientific Information in Horizon 2020

Celina Ramjoué

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

International Conference on Science 2.0
26-27 March 2014, Hamburg
The broader policy context:
Opening up scientific information is just one piece of the picture:
Open, Digital Science
Open, Digital Science

**Higher impact science:** impact on society and the economy

**ICT-enabled transformation of science:**
- New research methods and ways of evaluating research
- Opening up research results (scientific publications & research data)
- Open research collaborations, e-infrastructures
- Citizen engagement (citizen science, crowd-sourcing)

**More efficient science:** shared resources, dynamic collaborations

**New science:** New disciplines, new research topics

**Better science:** Transparent and replicable research
Why open up scientific information?

**Goal:** optimise the impact of publicly-funded research and innovation

**Expected impacts of opening up scientific information:**

- Better science (build on previous results)
- More efficient science (avoid duplication & promote re-use)
- Economic growth (accelerated and open innovation)
- Improved transparency (involving citizens & society)

**How?**

- Open up scientific information resulting from EU-funded research (Horizon 2020)
- Work with Member States to encourage co-ordination of policies

**Political basis:** Scientific information package (Communication & Recommendation to MS) and ERA Communication, July 2012
What scientific information?

1. Scientific publications:
   **Open Access (OA):** online access at no charge to the user
   Two main OA publishing business models
   - **Self-archiving:** deposit of manuscripts & immediate/delayed OA provided by author ("Green OA")
   - **OA publishing:** costs covered & immediate OA provided by publisher ("Gold OA")

2. Research Data:
   **Open Research Data (ORD):** data that can be accessed, mined exploited, reproduced and disseminated – free of charge for any user

**Scientific information: increasingly blurred boundaries**
- **Scientific publications ... are data**
  - Text is data (text and datamining)
  - Underlying research data
- **Research data can be published (data publications)**
Opening up scientific information in Horizon 2020 ...
Basis: OA to publications in FP7

**OA Pilot in FP7**
- "Best effort" basis
- 7 areas (>1800 projects to date)
- 20% of total FP7 budget (2007-2013)
- Support from researchers, need for more support

**OA (gold) publishing costs eligible in FP7**
- Since the beginning of FP7, for all projects
- Limited to duration of project

**e-Infrastructure: OpenAIRE / OpenAIRE+**
- EU-funded portal giving access to repositories across Europe (implements FP7 Pilot), network of helpdesks
- 47,000 publications, 19,000 OA, others still under embargo/restricted
OA to publications mandate: obligations

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

- Deposit a machine-readable copy of the published version or final peer-reviewed manuscript accepted for publication in a repository of the researchers choice (possibly OpenAIRE compliant)
- Ensure OA on publication or at the latest within 6 months (12 for SSH)
- **New**: Aim to deposit at the same time the research data needed to validate the results ("underlying data")
- **New**: Ensure OA to the bibliographic metadata that identify the deposited publication, via the repository
OA to publications mandate: other issues

Routes towards OA:

- OA publishing/gold and self-archiving/green considered valid and complementary routes
- Always deposit into a repository (also in the case of gold OA)

Costs for OA publishing:

- Eligibility of OA publishing costs during the grant (as in FP7)
- Piloting a mechanism for open access publishing after the end of the grant agreement (call EINFRA-2-2014 – eInfrastructure for Open Access)

Licencing:

- Encourage authors to retain their copyright and to grant adequate licences to publishers (e.g. Creative Commons)
New: Pilot on Open Research Data in Horizon 2020:

- Scope of the Pilot?
- When can actions opt out?
- What data is covered?
- What about data management?
- What are the obligations?
Pilot on Open Research Data: Scope

Areas of the 2014-2015 Work Programme participating in the Open Research Data Pilot are:

- Future and Emerging Technologies (FET)
- Research infrastructures – part e-Infrastructures
- Leadership in enabling and industrial technologies – Information and Communication Technologies (LEIT-ICT)
- Societal Challenge: Secure, Clean and Efficient Energy – part Smart cities and communities
- Societal Challenge: Climate Action, Environment, Resource Efficiency and Raw materials – except raw materials
- Societal Challenge: Europe in a changing world – inclusive, innovative and reflective Societies
- Science with and for Society

Actions in other areas can participate on a voluntary basis!
Pilot on Open Research Data: opting out

Actions 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
- In case of conflict with the obligation to protect results
- In case of conflict with confidentiality obligations
- In case of conflict with (national) security obligations
- In case of 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
Pilot on Open Research Data: What data?

**Types of data concerned:**

- Data (including associated metadata) needed to validate the results presented in scientific publications ("underlying data")
- Other data (including associated metadata) as specified in a data management plan (DMP)
What about data management?

- New focus on Data management in H2020
- All proposers to submit general information on data management - evaluated under criterion 'Impact'
- Data Management Plans (DMPs) mandatory for all actions participating in the Pilot (deliverable within the first six months)
- Other projects invited to submit a DMP if relevant for their planned research
- DMP questions (template: Data Management Guidelines):
  - What data will be collected or generated?
  - What standards will be used and how will metadata be generated?
  - What data will be exploited? What data will be shared /made open?
  - How will data be curated and reserved?
Beneficiaries participating in the Pilot will:

- Deposit a) underlying and b) "other data" as specified in the DMP into a research data repository of their choice
- Take measures to make it possible to access, mine, exploit, reproduce and disseminate free of charge (using e.g. Creative Commons licences)
- 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)
- **Note:** Actions participating in the Pilot are not obliged to make all datasets open (as described in their DMP, compliance with confidentiality, security, data protection, etc.)
ORD Pilot: a chance to co-shape policy

- Opening up research data: the new frontier
- Ambitious, yet pragmatic design of the pilot: droad scope, opt-out, voluntary participation possible
- Pilot is flexible; numerous safeguards in place
- Uptake of and experiences with the Pilot will be monitored
- Need to collect and analyse many and varied experiences
- Support & monitoring to be developed

- Participating in the Pilot means co-shape European policy on opening up research data ... in the next Framework Programme!
Opening up scientific information in Horizon 2020

Thank you!

celina.ramjoue@ec.europa.eu

http://ec.europa.eu/digital-agenda/en/open-access-scientific-knowledge-0
http://ec.europa.eu/research/science-society/open_access
