The state of open research data in Horizon 2020

Daniel Spichtinger
DG RTD, Unit A.6.

Daniel.spichtinger@ec.europa.eu

European Commission

Content

Background: political support and legal base

State of Play: provisions in H2020 and uptake

Current/Future action and challenges to be solved

The Need to be Open

Open Science

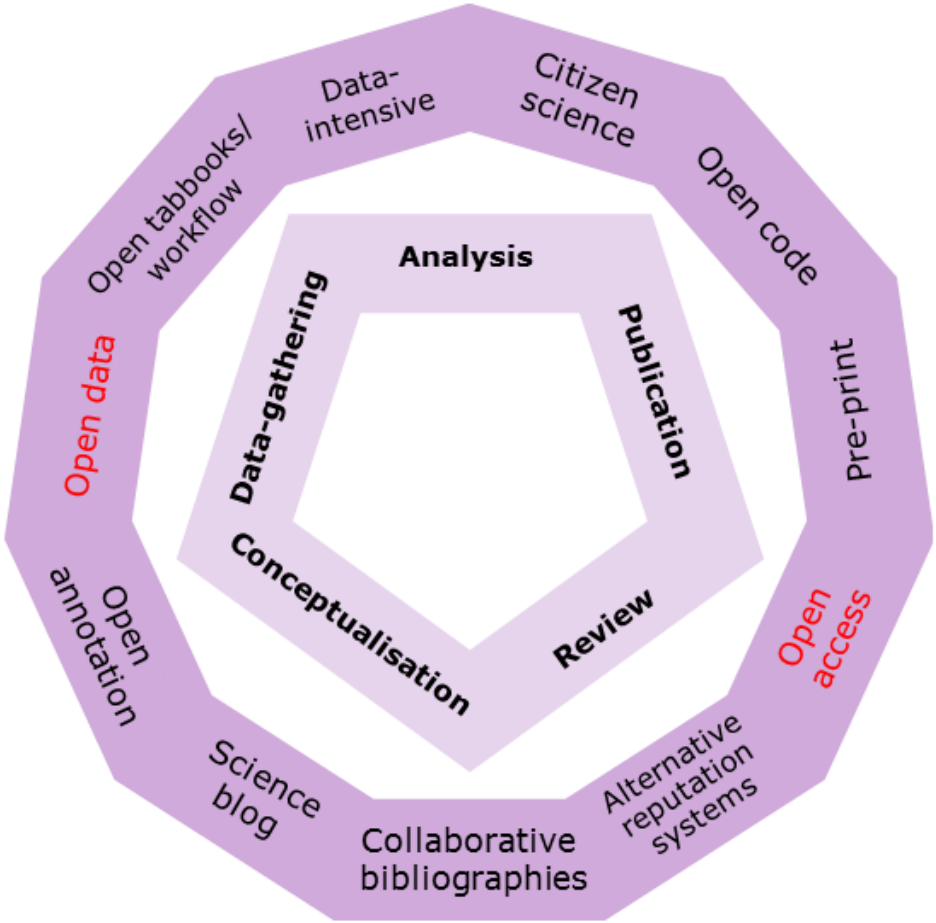
A systemic change in the modus operandi of science and research

Affecting the whole research cycle and its stakeholders

Commissioner Carlos Moedas
Open Science Presidency Conference
Amsterdam, 4 April 2016

European Commission

Open Science – opening up the research process



Feeding into the EC's Open Science policy

OA to publications & research data

Open access to publications and research data is a cornerstone of the Commission's broader Open Science policy

Extensive stakeholder consultation

- ✓ Public consultation (July-September 2014)
- ✓ Validation workshops (October-December 2014)
- ✓ Final report (February 2015): http://ec.europa.eu/research/consultations/science-2.0/science_2_0_final_report.pdf

Broad consensus on five policy action lines

- ✓ Policy debate & Council conclusions 'data-driven economy' May 2015
- ✓ Presidency conference Open Science & Council conclusions 'open science') May 2016

Reflected in the Commission top priorities and actions

- ✓ Included in the Digital Single Market strategy May 2015
- ✓ European Open Science Agenda May 2015
- ✓ High Level Expert Groups on 8 Action Lines

Open Science: key areas

1. Reward systems
2. Measuring quality and impact: altmetrics
3. Changing business models for publishing
4. FAIR open data
5. Open Science Cloud
6. Research integrity
7. Citizen Science
8. Open education and skills

Background

Member States

2012 Recommendation on Access to and Preservation of Scientific Information – National Points of Reference (NPR)

- 2016 report, 2017 update in progress (likely end of the year)

ERAC Taskforce on Open Data

Open Science Presidency Conference Amsterdam

Amsterdam Call for Action on Open Science

May 2016 Council Conclusions

- Support for OA publications, data re-use & data management

Stakeholders

Open Science Policy Platform

ERAC TF on Optimal Reuse of Research Data

TRAINING OF STAKEHOLDERS AND AWARENESS RAISING

1. Promote a better understanding of open research data

2. Establish training and education programs on Open Science

3. Establish a reward system for data sharing activities

4. Ensure sound monitoring

DATA QUALITY AND MANAGEMENT

5. Make data identifiable and citable

6. Promote metadata standardisation and production of metadata

7. Promote innovative models for peer-review and quality assurance

8. Strongly promote the use of data management plans

SUSTAINABILITY AND FUNDING

9. Ensure the existence of FAIR open research data infrastructures

10. Ensure funding for open research data and for data sharing activities

LEGAL ISSUES

11. Make IPR issues insightful

Amsterdam Call for Action

Two important pan-European goals for 2020:

1. **Full open access for all scientific publications**
2. **A fundamentally new approach towards optimal reuse of research data**

Flanking policy

3. **New assessment, reward and evaluation systems**
4. **Alignment of policies and exchange of best practices**

Political Support
Council Conclusions 27 May 2016

WELCOMES OA to scientific publications as the default. Target of 100% OA by 2020

SUPPORTS optimal re-use of data with the underlying principle of "as open as possible as closed as necessary"

WELCOMES the intention of the Commission to make research data produced by Horizon 2020 open by default, whilst recognising the right of opting out

CALLS on the Commission to promote data stewardship, including DMPs – importance of making data findable, accessible, interoperable and re-usable (FAIR)

Open access in Horizon 2020

Regulation establishing Horizon 2020

"To increase the circulation and exploitation of knowledge, open access to scientific publications should be ensured. Furthermore, open access to research data resulting from publicly funded research under Horizon 2020 should be promoted, taking into account constraints pertaining to privacy, national security and intellectual property rights"

*Open access to **scientific publications** resulting from publicly funded research under Horizon 2020 shall be **ensured** [...].*

*Open access to **research data** resulting from publicly funded research under Horizon 2020 shall be **promoted**. [...].*

Output-& Impact-Oriented Focus of Horizon 2020

Dissemination | **Exploitation** | **Communication**

open access

Why open access in Horizon?

***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

Open Research data

Open Research Data Pilot (2014-2016)

As of the Work Programme 2017: extended to cover all thematic areas of Horizon 2020 ('open by default')

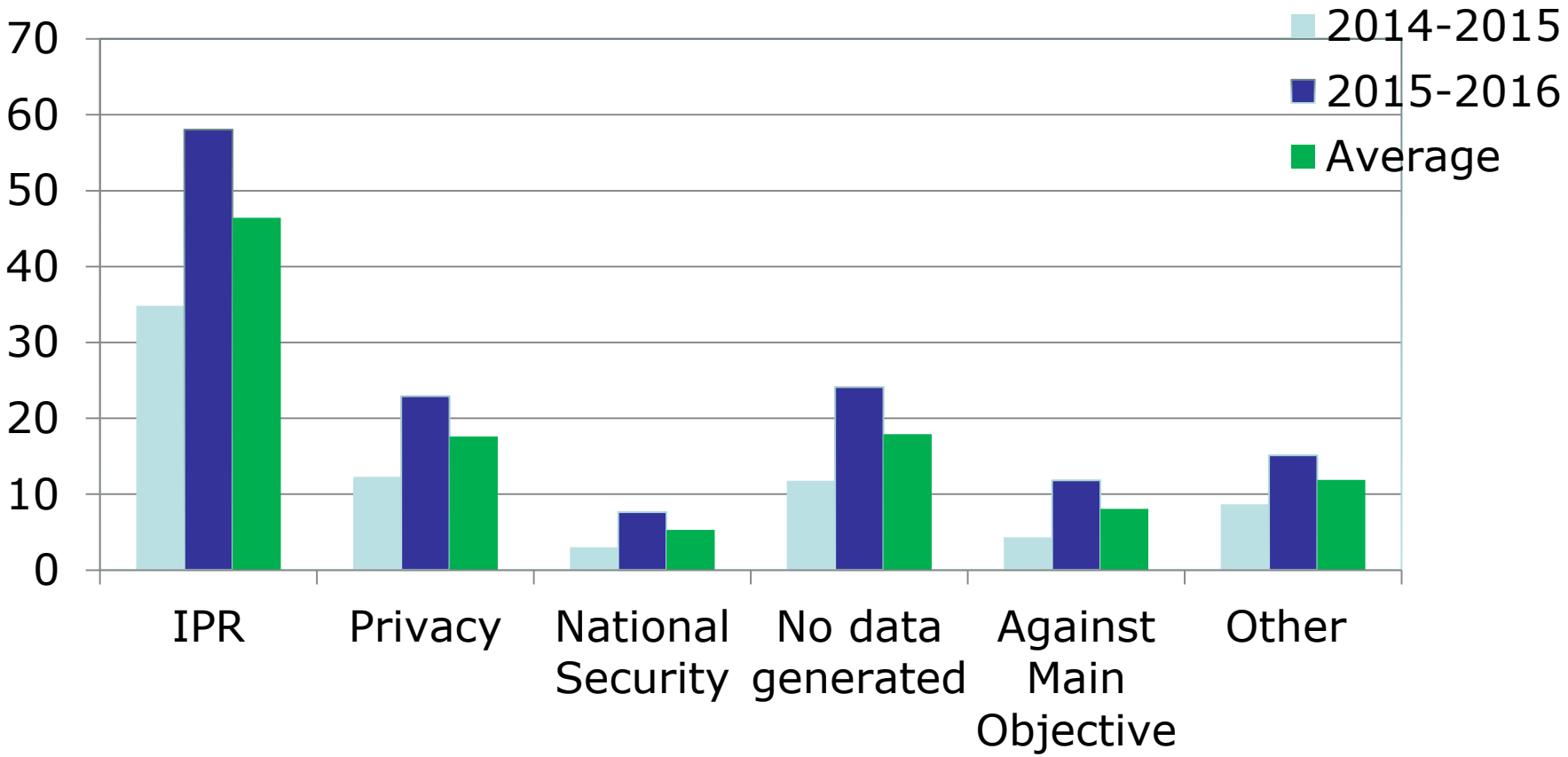
- Projects may still opt-out at any stage (IPR, personal data protection, national security, other reasons)
- Mainly concerns data underlying publications (other data optional)
- Data Management Plan obligatory by M6 (not part of project evaluation)

Approach: as open as possible, as closed as necessary

Results from the Open Research Data Pilot (Finalised data 2014-2016)

Year	Core areas (CA)		Non-core areas (NCA)	Sample Size (signed GAs)
---	Stay in	Opt-out	Opt-in	
2014-2015	67,44%	32,56%	2,5%	390 (CA) 3244 (NCA)
2015-2016	68,13%	31,87%	15,25%	935 SGA (CA) 6250 (NCA)
Average	**67,78%**	32,21%	8,87%	

Opt-out reasons



European Commission

RESEARCH DATA - OPEN BY DEFAULT

Clarifying terminology...

In the past our policy mainly addressed the 'accessibility' part of FAIR.

- Started off with 'open access to research data'
- Moved towards open (research) data with the ORD pilot (which also covered further aspects)
- We are now seeing openness as one component of FAIR data and aim to address all of the FAIR aspects in Horizon 2020

A FAIR DMP has to address that data are

- '**Findable**', i.e. discoverable with metadata, identifiable and locatable by means of a standard identification mechanism;
- '**Accessible**', i.e. always available and obtainable;
- '**Interoperable**', i.e. both syntactically parseable and semantically understandable, allowing data exchange and reuse between researchers, institutions, organisations or countries; and
- '**Reusable**', i.e. sufficiently described and shared with the least restrictive licences, allowing the widest reuse possible and the least cumbersome integration with other data sources.

FAIR Data Management DMP

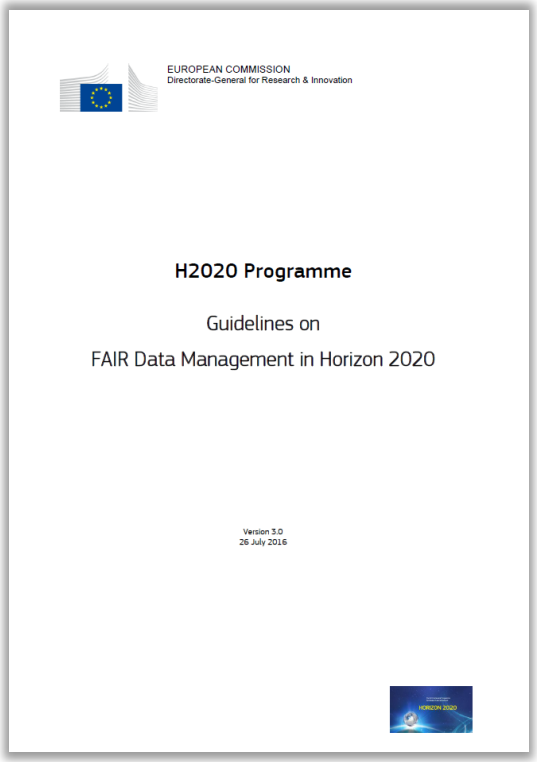
- Template DMP (Annex to Guidelines on FAIR Data Management)
  - Provided as a service, its use is currently optional
- Standard DMP template is light and flexible
  - Set of questions + summary table
- One DMP per project not per dataset
  - but mention if there are specific issues for a particular dataset)
- DMP as a living document
  - Updated as part of periodic evaluation and/or at least at the end of the project for final reporting

Guidelines on FAIR Data Management

Available here on the Participant Portal!

Read it!!!

Only 12 Pages!!!



Source: REA 2016 assessment of H2020 Societal Challenge 6 projects.

Initial DMP experiences

Additional guidance on data management is needed for all groups of actors in research projects (researchers, peer reviewers and funder administrators ('project officers') including roles supporting researchers with data management tasks (data librarians or IT professionals working in data centres).

Aspects such as data preservation, IPR or standards are too often not well developed in the DMPs that have been submitted so far

Nevertheless research projects with excellent RDM performance are not rare. Some high quality DMPs from H2020 projects have already been published online, see

http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

ORD Pilot: experiences (2)

- Explanation is paramount!
- Misperception that 'open' bias will be 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.

ORD Pilot: experiences (3)

- 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

Challenges to be solved & Tools

- **Open access to publications**

  Increasing uptake to 100% - reinforced monitoring and incentives

- **Open access to research data**

  Mainstreaming FAIR data across Horizon 2020 & FP9 – requires a change in scientific culture

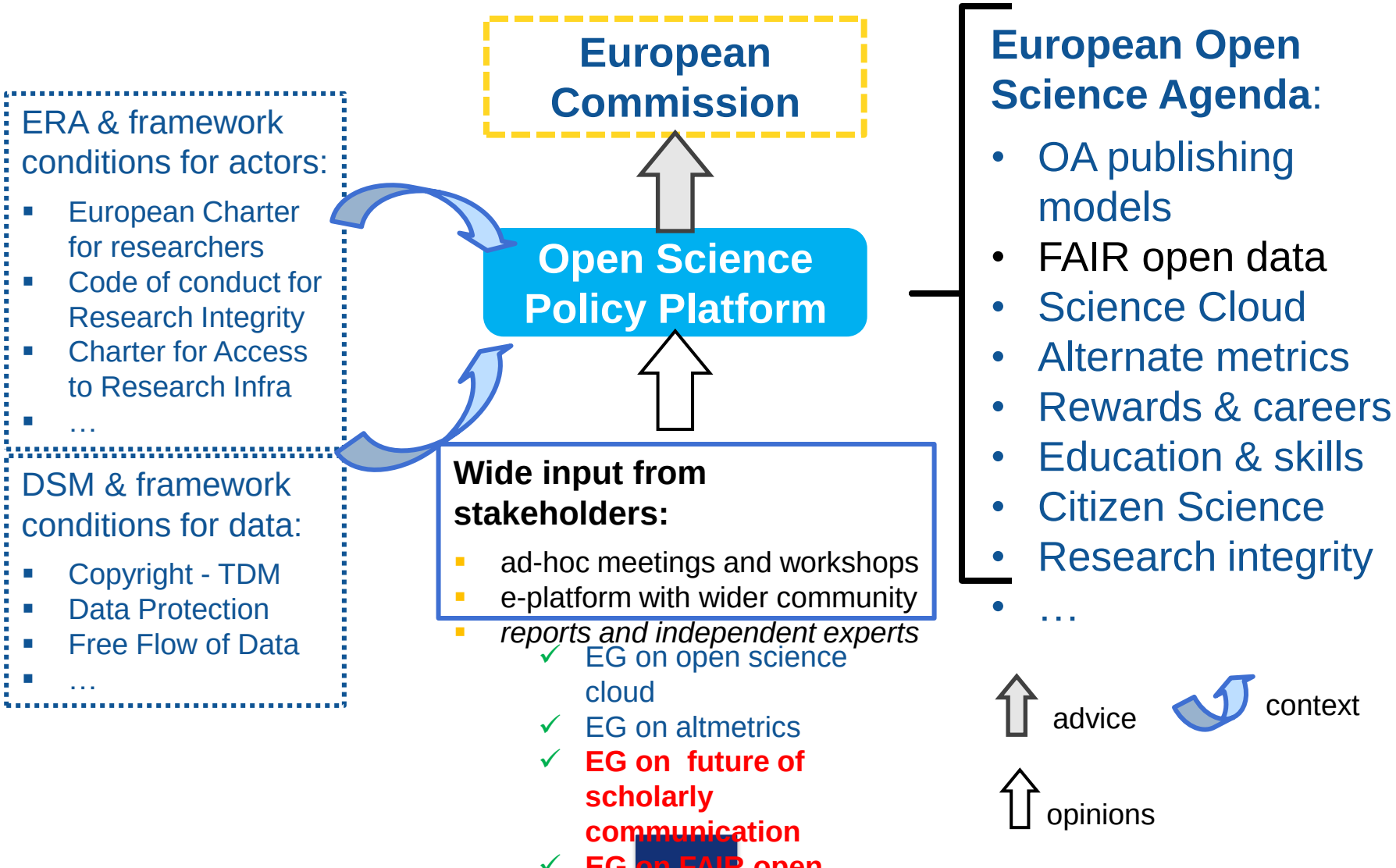
- **Tools to further develop policy**

  Open Science Policy Platform

  Commission Expert Groups on FAIR data & on Future of scholarly communication

  continuing dialogue & cooperation with Member States (RWP, ERAC, NPR, MLE)

Open Science Policy Platform



Ressources

Open innovation, open science, open to the world. A vision for Europe

http://bookshop.europa.eu/en/open-innovation-open-science-open-to-the-world-pbKI0416263/

NPR report

https://ec.europa.eu/research/openscience/pdf/openaccess/npr_report.pdf#view=fit&pagemode=none

H2020 guidance

http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

NEW May 2016 Council Conclusions

http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf

NEW Uptake of ORD pilot (dataset)

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

European Commission

We welcome your input

Contact us

Mail: RTD-open-access@ec.europa.eu

Web: http://ec.europa.eu/research/openscience/index.cfm

Twitter: @OpenAccessEC

Questions and Answers