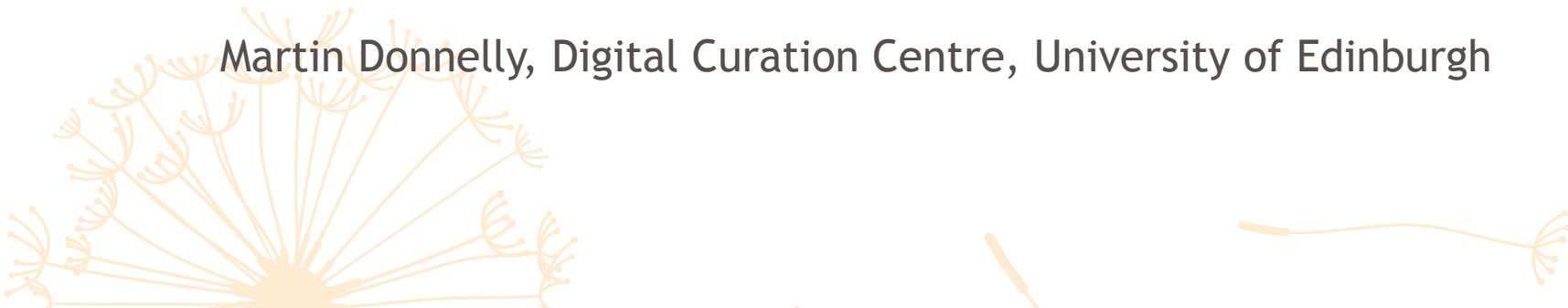




Facilitate Open Science Training for European Research

An overview of open science and open data in H2020

Martin Donnelly, Digital Curation Centre, University of Edinburgh





Facilitate Open Science Training for European Research



OVERVIEW

1. Open Access and Open Data
2. Open Access in FP7 and Horizon 2020
3. The Data Management Pilot in Horizon 2020
4. About the FOSTER project



1. OPEN ACCESS AND OPEN DATA

- Open Access (OA) is part of a broader trend in research, sometimes termed ‘Science 2.0’. The EC sees a real economic benefit to OA by supporting SMEs and NGOs that can’t afford subscriptions to the latest research. A study by [Houghton, Swan and Brown](#) provides quantifiable evidence of how much a lack of OA costs SMEs, both in terms of the time lost accessing documents and the delays to produce new products
- OA is past the ‘tipping point’ in several fields (e.g. biology, biomedical research, mathematics and general science & technology) whereas the social sciences, humanities, applied sciences, engineering and technology are the least engaged. (These findings based on a [Science Metrix study](#).)
- Some disciplines have committed to sharing data and are reaping the benefits - the research process is fastest in High Energy Physics now due to the community practice of immediate sharing
- The EC ran an OA pilot under FP7, which is now being followed up with a requirement under Horizon 2020. A pilot for open data has also been introduced with an intention to develop policy in the same way...



1a. Background: Open Access and RDM

- Open Access (OA) was born in the 1980s with free-to-access Listserv journals, but it really took off with the popularisation of the Internet in the mid-1990s, and the subsequent boom in online journals
- The Internet lowered (physical) barriers to accessing knowledge, but financial barriers remained - indeed, the cost of online journals tended to increase much faster than inflation, and scholars/libraries faced a cost crisis
- As Open Access to publications became normal (if not ubiquitous!), the scholarly community turned its attention to the data which underpins the research outputs, and eventually to consider it a first-class output in its own right. In fact, the development of the OA and research data management (RDM) agendas are closely linked...



Timeline: Open Access and Data Sharing

- 1987: New Horizons in Adult Education launched by the Syracuse University Kellogg Project. (An early free online peer-reviewed journal.)
- 1991: The “Bromley Principles” Regarding Full and Open Access to “Global Change” Data, in Policy Statements on Data Management for Global Change Research, U.S. Office of Science and Technology Policy
- 2001: The term “Open Access” (OA), the free online availability of research literature, is first coined at an Open Society sponsored meeting in Budapest, Hungary.
- 2004: Ministerial representatives from 34 nations to the Organisation for Economic Co-operation and Development (OECD) issue the Declaration on Access to Research Data From Public Funding.
- 2006: The Scientific Council of the European Research Council (ERC) pledges to adopt an OA mandate for ERC-funded research “as soon as pertinent repositories become operational”.
- 2012: European Commission recognises research data is as important as publications. Announces in July 2012 that it would experiment with open access to research data (see IP/12/790)

http://europa.eu/rapid/press-release_IP-12-790_en.htm

(Derived from, *inter alia*, Peter Suber (2009) “Timeline of the open access movement”, <http://www.earlham.edu/~peters/fos/timeline.htm>)



2a. Recap: Open Access in FP7

The EC's Open Access pilot ran from August 2008 until the end of the Seventh Research Framework Programme (FP7) in 2013. It required grant recipients in certain areas to "deposit peer reviewed research articles or final manuscripts resulting from their FP7 projects into an online repository and make their best efforts to ensure open access to these articles." Both green and gold OA were catered for.

- **Rationale:**
 - to improve and promote the dissemination of knowledge, thereby
 - improving the efficiency of scientific discovery, and
 - maximising return on investment in R&D by public research funding bodies
- **Coverage:** Peer reviewed research articles in the following areas...
 - Energy; Environment (including Climate Change); Health; Information and Communication Technologies (Cognitive Systems, Interaction, Robotics); Research Infrastructures (e-infrastructures); Science in society *; Socio-economic sciences and the humanities *
- **Timing:** Open access to these publications is to be ensured within six months after publication (* twelve months in the last two areas)
- **Place of deposit:** Institutional repository was first choice, failing that "an appropriate subject based/thematic repository" or the EC's open repository for papers that would otherwise be homeless.
- **Full guidelines:** ftp://ftp.cordis.europa.eu/pub/fp7/docs/open-access-pilot_en.pdf



2b. From Open Access to Open Science

- “The European Commission is now moving beyond open access towards the more inclusive area of open science. Elements of open science will gradually feed into the shaping of a policy for Responsible Research and Innovation and will contribute to the realisation of the European Research Area and the Innovation Union, the two main flagship initiatives for research and innovation.”

<http://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=science>

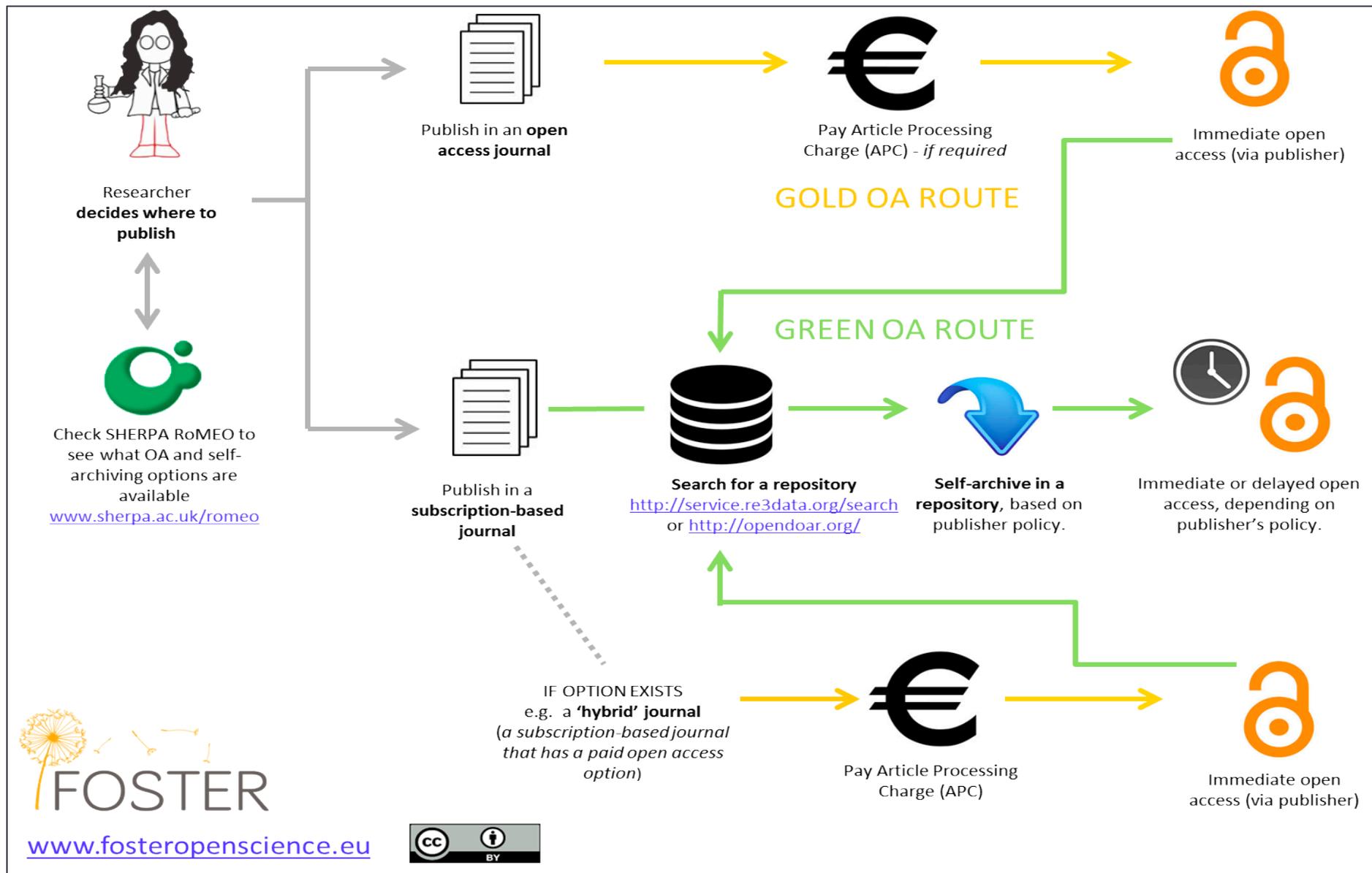
- All projects receiving Horizon 2020 funding are obliged to make sure that any peer-reviewed journal article they publish is openly accessible, free of charge.



2c. H2020 Publishing particulars

- The EC view is that the H2020 OA mandate does not seek to restrict publishing in any way. Researchers can publish where they choose. The only requirement is that they make sure the publication is made openly available via a repository. This can be done by:
 - publishing with an OA journal, which may or may not charge an APC;
 - publishing with a subscription-based journal, and depositing a copy into a repository (with open access being usually delayed by an embargo period imposed by the publisher); or
 - if the option is provided by the publisher, pay an APC to have an immediate open access copy.
- Under Horizon 2020, a copy of the article must always be deposited in a repository, even if the gold (or hybrid) has been chosen.
- When researchers first decide where to publish, it's useful to consult a service like SHERPA RoMEO to see what open access options are available. Researchers could start with a list of targeted journals and prioritise, or use a mix and match approach based on the results of this.
- Although over 60% of publishers don't charge APCs, fees can be quite steep. The average rate is €1,020 per article for open access publishers and €1,980 for hybrid journals. (Ref: [Björk & Solomon](#)). It could be very costly to always choose the gold route and pay lots of APCs, so a mixture of gold and green approaches is probably best.

2d. Possible OA pathways



2e. Summary points

- Main points of the Horizon 2020 Open Access requirements:
 - Researcher chooses where to publish;
 - Requirements apply to peer-reviewed articles rather than monographs, technical reports and conference proceedings, though these can be included as desired;
 - All peer-reviewed publications should be made OA via the [green or gold](#) routes;
 - It is no longer sufficient to make publications available on the project website. Deposit in repositories is required in ALL cases (even under gold OA), so the bibliographic data is open and can be harvested by services like OpenAIRE;
 - The EC does not currently impose any price cap on fees for publication costs. Researchers should plan OA from the proposal stage, and write any APCs into the proposal under the dissemination budget;
 - The EC recommends how their funding should be mentioned in publications. This style should be followed in order to facilitate indexing.
- The primary document to consult is:
[Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#)



3. DATA MANAGEMENT IN H2020



- FP7 featured an **Open Access pilot**, and this is now an across-the-board requirement in FP8 (Horizon 2020 / H2020)
- H2020 includes an **Open Research Data pilot**, and it seems likely that it will become an across-the-board requirement in FP9...
- The main goals of these developments are to **lower barriers to accessing** the products of publicly funded research (“science”), and to **strengthen the integrity and longevity** of the scholarly record
- This section of the presentation focuses on the data management aspects of the Open Research Data pilot...

3a. Recap: Data Sharing and Publication

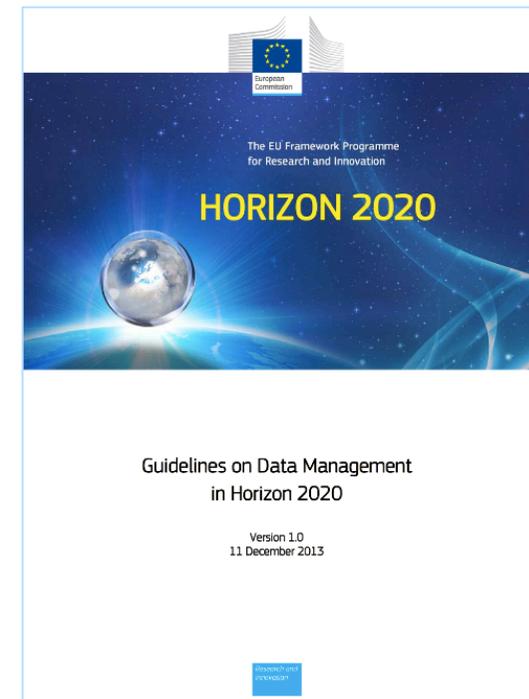
Benefits of sharing / publishing data...

- **TRANSPARENCY:** The evidence that underpins research can be made open for anyone to scrutinise, and attempt to replicate findings. This leads to a more robust scholarly record.
- **EFFICIENCY:** Data collection can be funded once, and used many times for a variety of purposes.
- **ACCESSIBILITY:** Interested third parties can (where appropriate) access and build upon publicly-funded research resources with minimal barriers to access.



3b. RDM in H2020: Overview

- The Horizon 2020 Open Research Data pilot covers “Innovation actions” and “Research and Innovation actions”
- It involves three iterations of Data Management Plan (DMP)
 - 6 months after start of project, mid-project review, end-of-project (final review)
- DMP contents
 - Data types; Standards used; Sharing/making available; Curation and preservation
- There are opt-out conditions. A detailed description and scope of the Open Research Data Pilot requirements is provided on the Participants’ Portal...



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

3c. Open Research Data Pilot: specifics (i)

AIM

The Open Research Data Pilot aims to improve and maximise access to and re-use of research data generated by projects. It will be monitored throughout Horizon 2020 with a view to further developing EC policy on open research.

SCOPE

For the 2014-2015 Work Programme, the areas of Horizon 2020 participating in the Open Research Data Pilot are:

- Future and Emerging Technologies; Research infrastructures; part e-Infrastructures; Leadership in enabling and industrial technologies; Information and Communication Technologies; 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

This corresponds to about €3 billion or 20% of the overall Horizon 2020 budget in 2014-2015.

COVERAGE

The Open Research Data Pilot applies to two types of data:

1. the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;
2. other data, including associated metadata, as specified and within the deadlines laid down in the data management plan.



3c. Open Research Data Pilot: specifics (ii)

STEP 1

- The data should be deposited, preferably in a dedicated research data repository. These may be subject-based/thematic, institutional or centralised.
- EC suggests the Registry of Research Data Repositories (www.re3data.org) and Databib (<http://databib.org>) for researchers looking to identify an appropriate repository
- Open Access Infrastructure for Research in Europe (OpenAIRE) will also become an entry point for linking publications to data.

STEP 2

- So far as possible, projects must then take measures to enable for third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) this research data.
- EC suggests attaching Creative Commons Licence (CC-BY or CC0) to the data deposited (<http://creativecommons.org/licenses/>, <http://creativecommons.org/about/cc0>).
- At the same time, projects should provide information via the chosen repository about tools and instruments at the disposal of the beneficiaries and necessary for validating the results, for instance specialised software or software code, algorithms, analysis protocols, etc. Where possible, they should provide the tools and instruments themselves.



3c. Open Research Data Pilot: specifics (iii)

COSTS

Costs relating to the implementation of the pilot will be eligible. Specific technical and professional support services will also be provided (e-Infrastructures WP), e.g. EUDAT and OpenAIRE, alongside support measures such as FOSTER.

OPT-OUTS

Opt outs are possible, either totally or partially. Projects may at any stage opt out of the Pilot for a variety of reasons, namely:

- if participation in the Pilot on Open Research Data is incompatible with the Horizon 2020 obligation to protect results if they can reasonably be expected to be commercially or industrially exploited;
- confidentiality (e.g. security issues, protection of personal data);
- if participation in the Pilot on Open Research Data would jeopardise the achievement of the main aim of the action;
- if the project will not generate / collect any research data;
- if there are other legitimate reasons to not take part in the Pilot (to be declared at proposal stage)



4. ABOUT THE FOSTER PROJECT



Facilitate Open Science Training for European Research





Facilitate Open Science Training for European Research

OBJECTIVES

- **Support different stakeholders**, especially young researchers, in adopting open access in the context of the European Research Area (ERA) and in **complying with the open access policies and rules of participation set out for Horizon 2020**;
- **Integrate open access principles and practice in the current research workflow** by targeting the young researcher training environment;
- **Strengthen institutional training capacity** to foster compliance with the open access policies of the ERA and Horizon 2020 (beyond the FOSTER project);
- **Facilitate the adoption, reinforcement and implementation of open access policies** from other European funders, in line with the EC's recommendation, in partnership with PASTEUR4OA project.





Facilitate Open Science Training for European Research

METHODS

- Identifying already existing **content** that can be reused in the context of the training activities and repackaging, reformatting them to be used within FOSTER, and develop/create/ enhance contents if/where they are needed.
- Creation of the **FOSTER Portal** to support **e-learning, blended learning, self-learning, dissemination of training materials/ contents and Helpdesk.**
- Delivery of **face-to-face training**, especially **training trainers/ multipliers** that can carry on further training and dissemination activities, within their institutions, countries or disciplinary communities.



THANK YOU

Martin Donnelly
Digital Curation Centre
University of Edinburgh

martin.donnelly@ed.ac.uk

Twitter: @mkdDCC

www.dcc.ac.uk

www.fosteropenscience.eu

