DATA MANAGEMENT PLANS IN HORIZON 2020

WHAT BENEFICIARIES THINK AND WHAT WE CAN LEARN FROM THEIR EXPERIENCE

DANIEL SPICHTINGER

INTEGRATIONG OPEN RESEARCH DATA INTO HORIZON 2020

- **2014: open research data pilot scheme (ORD Pilot)** in some thematic areas of the programme

- **As of Work Programme 2017: ORD Pilot extended to all thematic areas of Horizon 2020**

- But the Commission also recognises that there are good reasons to keep some or even all research data generated in a project closed ("opt-outs" e.g. due to personal data protection, IP)

- Main obligation: Create a Data Management Plan (DMP) by M6 – voluntary template from the EC available

- In recent years: focus on FAIR data (findable, accessible, interoperable, reusable)

(OPEN) RESEARCH DATA IN HORIZON EUROPE

- **"Evolution not revolution"**

- <mark>**Governing Principle to manage research data responsibly in line with FAIR and "as open as possible as closed as necessary"**</mark>

- At proposal stage applicants will be evaluated on preliminary RDM considerations, now included <mark>under "excellence"</mark> criteria

- All projects that generate/collect/re-use data will have to establish and regularly update a Data Management Plan

- Beneficiaries will have to deposit the data in a "<mark>trusted</mark>" repository

  - valid repositories will provide persistent identifiers & FAIR metadata

  - some actions may be obliged to deposit in a repository that is federated under the European Open Science Cloud

- Beneficiaries to ensure open access ASAP under CC-BY, CC-0 or equivalent unless exceptions apply (to be justified in the DMP)

- DMP to also include (information about) other research outputs (e.g. software, algorithms, protocols)

2. The DMP Use Case Project

The project

The aim of this project is to <mark>analyze Horizon 2020 DMP Use cases to identify good practices but also common challenges and mistakes amongst a number of use cases across different disciplines.</mark> The goal is to use these examples to support researchers with their DMP obligations throughout their own European projects. The results will also be reported in a publication.

The project was part of the OpenAIRE-Advance Projects RDM Task Force Group.

I am grateful to the IT team of the University Library of Vienna for their support.

This study has been supervised by Gerda McNeill (University of Vienna Library).

OpenAIRE

European Commission

universität wien
Universitätsbibliothek

Methodology

two main components to the project

- **qualitative part:**

  - ==qualitative analysis of 6 DMPs== (through modified rubric) and interviews with 6 cases studies (2 SH, 2 PE, 2 LS – ERC Classification, geographical and gender balance)

- **quantitative part**

  - Establishment of a while list of 840 DMPs downloaded from CORDIS, through a <u>manual</u> (is it s DMP / is it public) and <u>automated</u> screening process (screening for copyright) available at https://phaidra.univie.ac.at/detail/o:1140797, (content *not* screened)

  - ==survey of the DMP experiences of H2020 projects==

Survey Sample

- 108 respondents (Survey monkey)

- 87 gave project acronym (annex), 21 anonymous

- Questions could be skipped

- 63% ongoing, 37% completed (qualitative: 4 completed, 2 ongoing)

Results, Conclusions & Recommendations (RCR) form survey and interviews

Personal opinion of the author
not necessarily endorsed by Uni Wien and/or OpenAIRE

RCR according to themes

- Impact of the Mandate

- DMPs and project management

- Support

- Availability / curation of DMPS

RCR related to impact of the mandate-1

For **a significant number of projects (49/108) Horizon 2020 was the first time they encountered a Data Management Plan.** This underlines the importance of funder mandates to spread good data management practices.

Q3: Did you know about Data Management Plans before your Horizon 2020 project? n=108

ANSWER CHOICES	RESPONSES	
Yes	54.63%	59
No	45.37%	49
TOTAL		108

RCR related to impact-2

- Results from the qualitative interviews indicate that **knowledge increased over time** and that some have significantly developed their practices since their work on the project – this points towards the fast development of the area of data management but also to corresponding increase in competences.

- The **acceptance of data management plans is surprisingly high**: 82% find it useful or partly useful beyond it being an EC requirement

OpenAIRE   European Commission   universität wien Universitätsbibliothek

RCR related to impact-3

EC policy action / Horizon 2020 ORD mandate has had a significant impact

Q11: Do you consider the development of a data management plan useful beyond it being a requirement from the side of the European Commission

ANSWER CHOICES	RESPONSES	
Yes	53.27%	57
Partially	28.97%	31
No	17.76%	19
TOTAL		107

RCR related to project management-1

Having a DMP as part of the **Work Package on Management** (as opposed to dissemination or others) seems to be become the norm, in particular for small to mid-sized projects which do not have data science as their focus. We would therefore generally recommend projects to follow this approach, if there are no good reasons to do otherwise, but to also ensure links with the dissemination work package.

Q4: Which work package was/is the DMP part of?

ANSWER CHOICES	RESPONSES	
part of the project management work package	51.40%	55
part of the dissemination work package	21.50%	23
own work package for data management	17.76%	19
other, namely	9.35%	10
TOTAL		107

RCR related to project management-2

- In general, having **one person among each partner organization responsible for data issues** is a good practice (except potentially in very small projects). There also needs to be **one person that takes overall responsibility** for the project DMP – a DMP should not simply be made up of the parts delivered by the partners ("Frankenstein approach") but form an organic whole.

- For most respondents it was **neither very difficult nor very easy** to interact with partners on the DMP. On a scale of 1 (very easy) to 10 (very difficult), most answers were **clustered around 5-8**. Challenges encountered in the qualitative interviews related to **personal data** and GDPR, **amount of time and resources**, **coordination among geographically distant partner**s (akthough this is not limited to DM), type of data

- In one project **user agreements** were signed with the researchers – the data belongs to them but the project has limited usage rights

- One interviewee: data management does not necessarily mean open – some data was opened for scientific conferences (e.g. deposited on Zenodo)

RCR related to project management-3

- **Templates** are clearly important: 40% of the survey participants used the EC/ERC template. There were different opinions on the usefulness of templates, some ask for a more tailor made approach, which could be done by **providing a EC approved data management tool.**

Q7: Did you use a template or online tool when creating the data management plan for your project?

ANSWER CHOICES	RESPONSES	
YES- template from the European Commission/European Research Council	40.00%	42
YES - template from another organisation	17.14%	18
YES - online tool	8.57%	9
NO	25.71%	27
not applicable	8.57%	9
TOTAL		105

RCR related to support-1

- **Support** was primarily received from other partners, in some cases also the library and in a minority from OpenAIRE or from the IT department. 27% did not receive any support.

ANSWER CHOICES	RESPONSES	
YES- from other partners	39.42%	41
YES-from the library	11.54%	12
YES-from the IT departement	2.88%	3
YES- from OpenAIRE	3.85%	4
NO	26.92%	28
YES Other - please specify	15.38%	16
TOTAL		104

RCR related to support-2

- In the qualitative interviews, none of the participants received **content feedback** on the DMP from the EC but some did receive feedback from the reviewers. In the qualitative survey, the majority (56%) did not feedback either but those that did found it helpful

Q9: Was there feedback from the European Commission / the Agency you submitted the plan to?

ANSWER CHOICES	RESPONSES	
YES from the Project Officer	22.12%	23
YES from the Reviewer(s)	22.12%	23
NO	55.77%	58
TOTAL		104

RCR related to support-3

- Especially **beginners report a feeling of being lost** and, in particular before the template was available, had to do a significant amount of self-learning (qualitative interviews).

- A number of interviewees ask for a contact at the EC to contact for help

- **I would therefore recommend to set up a "one-stop-shop for Horizon research data management", akin to the IP helpdesk (could be done through a public procurement procedure)**

RCR related to availability/curation of DMPs-1

Interviewees often had to check whether their DMPs were available – and if yes, **most often made them available on their project website** (this chimes in with the finding from the quantitive survey) – however, several websites were no longer up and running

Q12: Did you publish your data management plan somewhere

ANSWER CHOICES	RESPONSES	
YES - in a repository	22.86%	24
YES - on the project website	30.48%	32
YES - somewhere else	8.57%	9
NO	38.10%	40
TOTAL		105

RCR related to availability/curation of DMPs-2

- This points to the importance of CORDIS as a source for public DMP which, however, is also not well known as a source for DMPs and thus also for the **need to raise awareness to deposit DMPs in (certified) repositories** to ensure preservation.

Summary

- The Horizon 2020 RDM requirement had a significant impact: many respondents (45%) first developed a DMP in H2020; knowledge about DM(P)s increased over the duration of H2020; 82% find DMPs useful or partially useful

- H2020 DMPs are mostly undertaken as part of the WP on management (51%), 40% used the EC/ERC templates – with mixed feelings. Support was mostly received from other project partners but 27% did not receive any support. Most participants (56%) did not receive feedback from the EC/REA. Beginners in particular tend to feel lost – these figures illustrate the need for a "DM Helpdesk"

- Most respondents did not publish their DMPs (38%) , those that did mostly put them on their website (30%)

More information & future plans

- Supplementary material on Phaidra, repository of the University of Vienna: https://phaidra.univie.ac.at/detail/o:1165751

- Publication in Open Research Europe: https://open-research-europe.ec.europa.eu/articles/1-42

  - (open) peer reviewed, comments to be implemented

- Blog: Data Management Plans in Horizon 2020 https://open-research-europe.ec.europa.eu/blog/data-management-plans-in-horizon-2020

- Upcoming: Analysis of the use creative commons licenses in DMPs (pending, draft submitted)

Thank you!

Daniel Spichtinger

daniel@spichtinger.net   danielsp   https://www.linkedin.com
/in/dspichtinger/