

Monitoring the Open Access Policy of Horizon 2020

Data Management Plan

As created for the study:

“Monitoring the Open Access Policy of Horizon 2020” (SPECIFIC CONTRACT No RTD/2019/SC/021 implementing Framework contract No 2018/RTD/A2/OP/PP-07001-2018)



PPMI, Athena Research Center, and UNU-MERIT



MOAP-Horizon2020-DMP¹

Funder

European Commission|EC

Grant

SPECIFIC CONTRACT No RTD/2019/SC/021 implementing Framework contract No 2018/RTD/A2/OP/PP-07001-2018

Organisations

- Athena Research and Innovation Center In Information Communication & Knowledge Technologies
- Public Policy and Management Institute
- United Nations University – Maastricht Economic and Social Research Institute on Innovation and Technology

Database DOI

- 10.5281/zenodo.4899767

Database documentation DOI

- 10.5281/zenodo.4900100

1. Database Summary

1.1 *What is the purpose of the data collection/generation and its relation to the objectives of the project?*

[To make informed decisions, To combine with other data]

Comment: This study is framed in the context of the contract entitled 'Monitoring the open access Policy of Horizon 2020 – RTD/2019/SC/021' reporting on a set of authoritative metrics on compliance to the European Commission open access mandate thus far in the Framework Programme, as well as advice on how to systematically monitor its compliance in the future.

1.2 *What types of data will the project generate?*

[derived or compiled (e.g., text mining, 3D models)]

1.3 *What formats of data will the project generate?*

[Text files]

The study produced CSV formatted datasets.

1.4 *What is the expected size of the data?*

Comment: The expected size of data is 1.8 GB.

1.5 *To whom might it be useful ('data utility')?*

[Research communities, Decision makers]

2. FAIR Data

2.1 *Making data findable, including provisions for metadata*

¹ This data management plan was created using Argos <https://argos.openaire.eu/splash/> .

2.1.1 Will you use metadata to describe the data?

Yes, Dublin Core.

Comment: The metadata standard used is OP Core from the Publications Office of the EC, which was developed according to the Dublin Core standard.

2.1.2 Will your metadata use standardised vocabularies?

Yes.

Comment: The standard vocabularies used are EuroSciVoc from the EU Vocabularies and the Frascati Manual.

2.1.3 Will you make the metadata available free-of-charge?

Yes

2.1.4 Will your metadata be harvestable?

Yes

2.1.5 Will you use naming conventions for your data?

Yes

Comment: We use one directory per database schema, with the name of the directory being the name of the schema. In each directory we store each schema table as a CSV file, with the name of the file being the name of the table. Example: moap/[schema name]/[table name].csv .

2.1.6 Will you provide clear version numbers for your data?

Yes

Comment: Although no new version of the data is anticipated after the end of the study, it is possible to have versioned datasets if needed.

2.1.7 Will you provide persistent identifiers for your data?

Yes, we will provide a DOI.

2.1.8 Will you provide searchable metadata for your data?

Yes, via public repositories.

2.1.9 Are the file formats you will use open?

Yes

2.2 *Making data openly accessible*

2.2.1 Are there ethical or legal issues that can impact sharing the data?

No

Comment: The derived dataset does not contain personal or confidential data.

2.2.2 Will your data be openly accessible?

Yes, all data will be shared publicly with open licenses so that they are reused and further exploited by others.

2.2.3 How will the data be made available?

[Repository of Archive]

Zenodo will preserve the data. Data might also be available in institutional repositories of the partners.

2.2.4 Is the storage sufficiently secure for the data and does the storage provide backup and recovery procedures?

Secure with backup and recovery

2.3 *Making data interoperable*

This dataset concerns metadata records and is provided in machine readable formats, according to community standards.

2.4 *Increase data reuse*

2.4.1. What internationally recognised licence will you use for your data?

Creative Commons Attribution 4.0 International

2.5 *How do you ensure data re-use?*

Via institutional archive.

3 Allocation of resources

3.1 *How will the cost of making your data findable, accessible, interoperable and reusable be covered?*

[Use of institution infrastructure]

Comment: Part of the costs will be covered by use of public infrastructures like Zenodo.

3.2 *Will you identify a data manager to manage your data, if not who will be responsible for the management of your data?*

Yes

Comment: Antonis Lempesis from the Athena Research Manager is the data manager.

4 Ethical aspects

4.1 *Are there any ethical or legal issues that can have an impact on data sharing?*

No

Datasets

The study utilizes two existing datasets as input to provide added value outputs regarding the monitoring of Open Access policies. Input datasets are: [moap_ec](#) and [moap_org](#). Output dataset is: [moap_final](#).

Output dataset

Dataset Title: moap_final

Dataset Description

This dataset contains the indicator tables as produced in the MOAP Horizon 2020 study.

Input – Reused datasets

Title: moap_ec

Dataset Description: This dataset contains data as reported in the European Commission's System for Grant Management related to Horizon 2020 funded publications and datasets.

1 Reusable Data

1.1 *Will you re-use any existing data and how?*

Yes

1.2 *Where do the data reside?*

Data have been shared by the European Commission.

1.3 *Which data will be re-used?*

All received data was used for the purpose of the study.

Title: moap_org

Dataset Description: This dataset contains data from OpenAIRE related to Horizon 2020 funded publications and datasets.

1 Reusable Data

1.1 *Will you re-use any existing data and how?*

Yes

1.2 *Where do the data reside?*

Data were acquired from OpenAIRE public API.

1.3 *Which data will be re-used?*

Data used in the study form a subset of Horizon 2020 data contained in the OpenAIRE Graph.