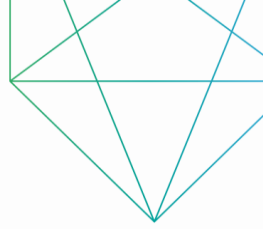


PUBLIC

Deliverable 6.6 Data Management Plan

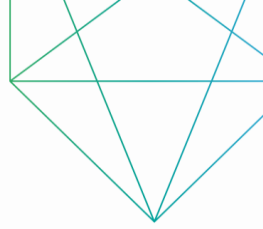
DR. ERIN SCHOLS

LINQ CONSULTING & MANAGEMENT



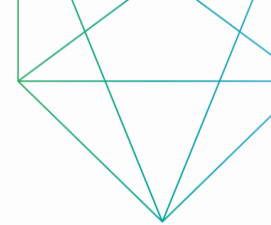
Deliverable information	
Authors	Erin Schols
Lead partner	Linq Consulting & Management
Relevant task	Task 6.4
Version number	1
Version date	29 June 2023

Dissemination level	
X	PU – Public (fully open; automatically posted online)
	SEN – Sensitive (limited under the conditions of the Grant Agreement)



Contents

List of tables	4
List of acronyms	4
1 Introduction	4
1.1 About this document	5
1.2 Key principles.....	5
2 Data Summary	6
3 FAIR DATA.....	8
3.1 Making data findable, including provisions for metadata	8
3.2 Making data accessible.....	9
3.2.1 Repository.....	9
3.2.2 Data	9
3.3 Making data interoperable.....	9
3.4 Increase data re-use	10
4 Other research outputs.....	10
5 Allocation of resources	10
5.1 Costs for making data FAIR	10
5.2 Allocation of Responsibilities.....	11
6 Data security	12
7 Ethics	12
8 Other issues.....	12



List of tables

Table 1 Data summary	7
Table 2 List of designated Partner Data Managers	11

List of acronyms

DMP	Data Management Plan
DOI	Digital Object Identifiers
FAIR	Findable, Accessible, Interoperable and Reusable
FTP	File Transfer Protocol
GDPR	General Data Protection Regulation
ICT	Information and Communication Technologies
PDM	Partner Data Manager

1 Introduction

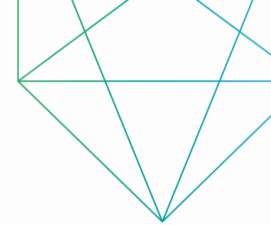
The Data Management Plan (DMP) of the OUTFOX project outlines how data emerging from the project will be handled during the lifetime of the project and after it is completed.

Horizon Europe seeks to ensure beneficiaries make their research data findable, accessible, interoperable and reusable (FAIR), to ensure it is soundly managed and exploited as effectively as possible. Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.

Task 6.4 Open Science & Research Data Management (duration M1-M48) is led by Linq Consulting & Management (LINQ) with all partners participating. LINQ will be the lead in the development and population of the Data Management Plan (DMP). The plan will map out the data that will be generated and held during the project by each partner. DMPs are a key element of good data management within collaborative research projects and help ensure the relevant requirements from the Grant Agreement are fulfilled.

The DMP describes the data management life cycle for the data to be collected, processed and/or generated within the project. As part of making research data FAIR, a DMP should include information on:

- The handling of research data during and after the end of the project
- What data will be collected, processed and/or generated
- Which methodology and standards will be applied



- Whether data will be shared/made open access
- How data will be curated and preserved (including after the end of the project).

1.1 About this document

This document (D6.6), based on the Horizon Europe template, provides a first version for the DMP that will be regularly updated and reviewed as a living document throughout the lifetime of the project. The document reviews and updates will take place at least every 9 months to reflect the status of the OUTFOX project with respect to its data management.

This document, along with the DMP appendix, seeks to capture and describe what data is collected, processed or generated and following what methodology and standards. The sections below also set out whether and how this data is shared and/or made open, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.

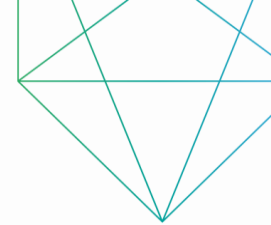
The OUTFOX project plans to collect, generate and manage data from and in the following countries:

- Netherlands
- Finland
- Italy
- Estonia
- United Kingdom

The required information to populate and update the DMP will be collected from partners using an accompanying questionnaire and updated, at a minimum, in time with the periodic evaluations/assessments of the project. In addition, workshops will take place during face-to-face consortium meetings as needed.

1.2 Key principles

- This DMP details what data the project will generate and manage, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.
- As part of the signed Consortium Agreement, ownership of key prior and post knowledge (IPR, data etc.) of all involved parties (including partners and cities) has been defined.
- The consortium will deposit all applicable data in a data repository, setup and maintained by LINQ (WP Leader and host of the project website). LINQ will take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate the following categories of data:
 - The data, including associated metadata, needed to validate the results presented in scientific publications;
 - Data collected during the project, after anonymization and including associated metadata, as specified in the DMP;



- Data generated during the project, including associated metadata, as specified in the consortium agreement and in the DMP;
 - Public project reports and public deliverables;
 - All material related to dissemination and communication actions.
- All data will be made available in XML format and/or in text/CSV (comma separated) format.
 - Data in databases will be accompanied by an open schema and fully documented specifications to allow for full and unrestricted accessibility.
 - The OUTFOX consortium will conform to the DMP guidelines and will make every effort through the project's dissemination activities to make this content findable, accessible, intelligible, and reusable (FAIR) by all interested stakeholders (especially by other related EU projects).
 - A searchable index and an information classification system of all relevant datasets will be developed, applicable to all cases and data types, while at the same time respecting all relevant IPR and copyright requirements pertaining to these datasets.
 - The initial DMP is delivered in Month 6 of the project. More elaborated and updated versions of the DMP will be made available around M18, M36, and M48, to fine-tune it according to the data generated and to the data uses that will be identified by the consortium.

2 Data Summary

A table of the types and formats of data is given below. The project will mainly generate new data, albeit based on previous experiments and knowhow of the project partners. This is a result of the project activities mostly involving new technologies and systems that will be developed and implemented within the project. In the case of benchmarking with existing technologies, the data will still be generated within the project to ensure consistency of the experimental conditions and replicability.

The tasks related to the techno-economic and life cycle assessments will re-use data from literature and database sources when the specific required values are either not generated in the project, or are generated at a later time than when needed for the assessments (with the assessments then updated as the project generated data becomes available).

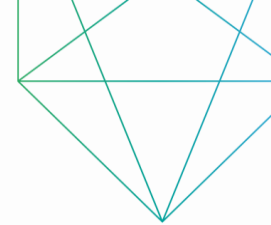
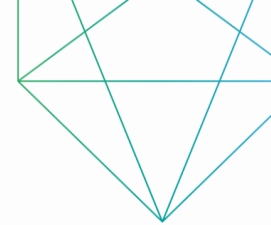


Table 1 Data summary

Type	Format	Purpose	Expected size	Origin	Data utility
Experimental data	Spreadsheets (.xls/.xlsx, .csv) Documents (.txt, .doc/.docx, PDF)	Evaluate experimental developments	10 GB	Generated	Share with the scientific community
Laboratory setup details & configurations	Spreadsheets (.xls/.xlsx, .csv) Documents (.txt, .doc/.docx, PDF)	Define and replicate experimental conditions	100 MB	Generated	Share with the scientific community and exploitation partners
Materials used	Spreadsheets (.xls/.xlsx, .csv) Documents (.txt, .doc/.docx, PDF)	Define and replicate experimental conditions	100 MB	Generated	Share with the scientific community and exploitation partners
Workshop results	Documents (.txt, .doc/.docx, PDF)	Process and evaluate input	< 1 MB	Generated	Share with policymakers and key stakeholders
Economic & market data	Spreadsheets (.xls/.xlsx, .csv) Documents (.txt, .doc/.docx, PDF)	Evaluate project developments	10 MB	Re-used & generated	Share with policymakers and key stakeholders
Environmental impact data	Spreadsheets (.xls/.xlsx, .csv) Documents (.txt, .doc/.docx, PDF)	Evaluate project developments	10 MB	Re-used & generated	Share with policymakers and key stakeholders
Photos	Image files (.jpg, .png)	Provide visual evidence and evaluation of project developments	1 GB	Generated	Share with key stakeholders
Videos	Video files (.mpeg)	Provide visual evaluation of project developments and means of communication	10 GB	Generated	Share with general public and key stakeholders



3 FAIR DATA

The following section outlines key elements that are to be considered in respect to the data collected and generated within OUTFOX. In general terms, OUTFOX research data should be “F.A.I.R.” that is

- Findable.
- Accessible.
- Interoperable.
- Re-usable.

These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation solution. It should be noted that participating in Open Science does not necessarily imply opening-up all OUTFOX research data. Rather, Open Science follows the principle “as open as possible, as closed as necessary” and focuses on encouraging sound data management as an essential part of research best practices. The European Commission recognizes that there are good reasons to keep some or even all research data generated in a project closed.

The DMP explains which data can be shared and under what terms and conditions, clearly separating legal and contractual reasons from voluntary restrictions. In a multi-beneficiary project like OUTFOX, it is also possible for specific beneficiaries to keep their data closed, if relevant provisions are made in the Consortium Agreement and are in line with the valid reasons for opting out of data sharing.

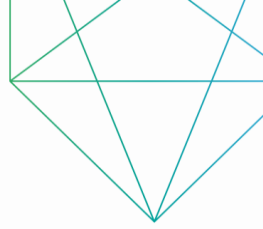
3.1 Making data findable, including provisions for metadata

Metadata is structured information which supports discovery, re-use and long-term storage of the data. There are three main types of metadata: descriptive, administrative, and structural.

- **Descriptive metadata** enables discovery, identification, and selection of resources. It can include elements such as title, author, and subjects.
- **Administrative metadata** facilitates the management of resources. It can include elements such as technical, preservation, rights, and use.
- **Structural metadata**, generally used in machine processing, describes relationships among various parts of a resource, such as chapters in a book.

The following points give a preliminary assessment of the metadata expected within the project and will be discussed in detail with the partners as the project progresses. Should updates be necessary, this document will be updated and a new version published.

- Datasets are to be in English.
- The following format is proposed for the naming of the files/documents:
Date(yyyymmdd)_OUTFOX_Partner_Filename_Versionnumber (e.g.
20230730_OUTFOX_TNO_D1.10 Next generation cells report_V0.1.docx)
- The metadata will contain the following information along with the datasets:
 - Keywords to ease the discovery and potential re-use of the data



- Creator of the data (Project partner and the responsible person)
- Data collection period
- Methods of data collection (e.g. source)
- Additional information on the abbreviations used
- Confidentiality information and dissemination rules of the metadata

3.2 Making data accessible

3.2.1 Repository

OUTFOX will use the following open repositories: i) Zenodo; ii) OUTFOX project website; and iii) arXiv for publications preprints. Zenodo is a well-known repository that fulfils the EU expectations for data accessibility, including secure long-term storage, citability, and versioning.

To ensure the safety of the data, TNO will use its local file servers and periodically create backups. All other relevant documentation created during the project will be archived and preserved in the OUTFOX SharePoint repository.

3.2.2 Data

The only data which will not be made openly accessible will be data which contains personally identifiable information (e.g. individual evaluation forms), data that supports deliverables that are kept confidential, and data identified as being commercially sensitive. The personal data processed in the project are not made publicly accessible but kept closed and inaccessible to third parties.

All data sharing will follow the provisions set out in the project Consortium Agreement. Furthermore, the generated data and deliverables are reviewed internally by a another technical expert before dissemination.

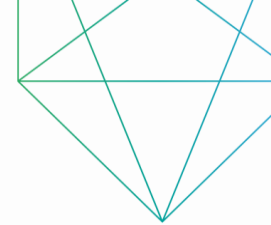
Data will be published using standard file formats (txt, pdf, csv etc.). All data will be accessible using standard tools. Software required to access the data should be made available, but it is not seen as being a requirement. Should it be needed, we will provide the required open source to access and analyze the data.

3.3 Making data interoperable

The data produced within OUTFOX are to be made interoperable by converting them to non-proprietary open formats or proprietary widely used formats, compliant with available software applications and facilitating the circulation and critical analyses among the partnership. In addition, harmonisation of the presentation of data (when disclosed in accordance with the project GA and CA) will be completed in order to maximise their impact. This will be an ongoing activity as the project progresses.

The vocabulary will be common for the same type of data in accordance with international regulation on Units/ Methods/ Regulations and harmonised among the partners using the format established by LINQ and agreed upon by all of the partners.

These vocabularies and formats will be established at the next General Assembly Meeting.



3.4 Increase data re-use

The exploitation of the data, methodologies and patents will be regulated according to the Consortium Agreement and included IPR rules. Also, the data disclosure to Third Parties will be regulated by a standard NDA, which will also be agreed upon by the consortium.

Open science data principles will be followed where possible. Supplementary research data will be made available in standard international formats with publications on open-access archives (e.g., arXiv.org) to enable reuse by other researchers. Researcher contact details will also be published with results presenting an avenue of last resort should researchers outside the consortium have questions about research methods, protocols and results. Commercially exploitable data (e.g., equipment and product designs) will be protected, and only made available by license.

The timetable on the data availability will be established according to the project GA, CA and NDA, all agreed upon and signed by each consortium partner. Also, the strategy of exploitation of the data after the end of the project will be agreed upon by the whole OUTFOX partnership, even if proposed by a specific committee.

TNO is in charge of the Quality Assurance Procedures, described in detail in Deliverable 7.2, and LINQ will support to further define the procedures related to data management, providing all data on samples, methodologies, and technologies, as requested.

4 Other research outputs

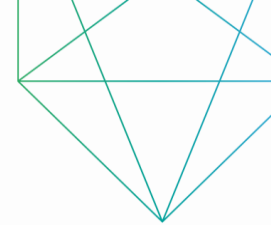
The other expected research outputs (other than data sets) include digital videos, news articles and newsletter to disseminate the project outputs, such as the demonstration of the OUTFOX technology. All dissemination materials will be shared through the project's communication channels (mailing list, social media, etc.) as well as being made available on the project website.

For each additional type of output, a data management strategy will be defined and agreed upon by the partners.

5 Allocation of resources

5.1 Costs for making data FAIR

No costs are foreseen to share the data on the OUTFOX SharePoint and Zenodo repository. Eventual costs related to data dissemination and publication are included in the project budget and eligible according the GA. The costs of dissemination will be related to the type and number of their disclosure in agreement with GA and IPR agreements in the CA.



5.2 Allocation of Responsibilities

People/groups involved in Data Management for OUTFOX are:

- The project coordinator (Cahit Benel, TNO)
- Research staff designing research, collecting, processing and analyzing data
- Laboratory or technical staff generating metadata and documentation
- External contractors involved in data collection, data entry, transcribing, processing or analysis
- Support staff managing and administering research and research funding, providing ethical review and assessing Intellectual Property Rights
- Institutional ICT services staff providing data storage, security and backup services (TNO, ICT departments)
- External data centers or web archives that facilitate data sharing (e.g. Zenodo)

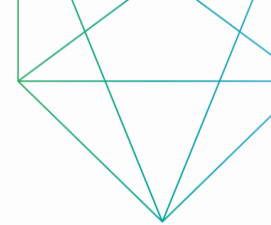
Each partner will designate a person ("Partner Data Manager", PDM), who will be responsible for any process or communication related to project data. All designated PDMs are listed in the table below:

Table 2 List of designated Partner Data Managers

Partner	Designated PDM	Email
TNO	Cahit Benel, supported by the Data Protection Officer at TNO	cahit.benel@tno.nl
VTT	Santeri Saxelin	santeri.saxelin@vtt.fi
Politecnico di Milano	Paolo Colbertaldo	paolo.colbertaldo@polimi.it
Fondazione Politecnico di Milano	Fabrizio Amarilli	fabrizio.amarilli@polimi.it
Elcogen AS	Sergii Pylypko	sergii.pylypko@elcogen.com
Elcogen OY	Antonio Alfano	antonio.alfano@elcogen.com
Convion	Tuomas Hakala	tuomas.hakala@convion.fi
Shell Global Solutions International	Srikanth Santhanam	Srikanth.Santhanam@shell.com
Linq Consulting & Management	Erin Schols	erin.schols@linq-consulting.com

The PDM is responsible on behalf of the corresponding partner for all data generated, collected, uploaded, processed, downloaded and stored in the partner's resources, or handled by that partner's personnel. Their responsibility extends to ensuring the following:

- That the datasets are accurate, consistent and complete.
- That the datasets are adequately documented.
- That the datasets are properly anonymized and their handling procedures are consistent with the GDPR and its related data management policies.
- Communication with LINQ for all issues related to project data and the DMP.



6 Data security

Data storage, in the context of Data Security, must be done in such a way to ensure the privacy and integrity of data and prevent unauthorised access, changes to data, disclosure or destruction of data. During the OUTFOX project, generated data intended for sharing between the project partners will be stored in a secure repository in Microsoft's SharePoint site, which will be managed by TNO. The access to the SharePoint site is possible using a TNO or a TNO Partner account, which is based on Multi-Factor Authentication. Microsoft offers standard facilities to ensure the confidentiality, integrity and availability of the information. In addition, information is backed-up on a daily basis.

Transmitting (uploading or downloading) sensitive or personal data between locations or within research teams must always be done using data encryption, e.g. using secure FTP or other secure data transfer protocol, to ensure data privacy and prevent unauthorized access of data (e.g. eavesdropping).

Access to data repositories will be password protected and access logs should be maintained.

Archived data of personal or sensitive nature should be stored encrypted, with strong encryption.

Before the OUTFOX project is completed, the partners will decide which data will have to be destroyed and which data will be maintained (and for how long). Retention time for curated datasets is by default twenty years.

To ensure data integrity, avoid loss of data and maintain storage consistency, regular data backups will be performed on a daily, weekly and monthly basis, either incremental or full. Data backups should be accompanied with appropriate and corresponding data recovery procedures.

7 Ethics

No ethical issues are foreseen for data collected or generated within OUTFOX.

8 Other issues

At this time, it is not foreseen that the project will make use of other national / funder / sectorial / departmental procedures for data management.