EERAdata Deliverable D30

Data Management Plan

Version 1 - Final

This data management plan establishes guidelines to manage all data collected, created, processed and published in the H2020 project EERAdata.

AUG. 2020

TOWARDS A FAIR AND OPEN DATA ECOSYSTEM IN THE LOW CARBON ENERGY RESEARCH COMMUNITY

Document description

D30. Data Management Plan

Work Package	Number, Name
Contributing organizations	HVL, IUE, AIT, ENEA, EERA, GIG
Preparation date	2020-08-25 (Month 6)
Due date	2020-08-31 (Month 6)
Nature of deliverable	report
Dissemination level	public
Version	No. 1
Dataverse identifier	XXX

Preparation slip

	Name	Organization	Date
From	August Wierling	HVL	23.08.2020
Edited by	Valeria Jana Schwanitz	HVL	25.08.2020
Reviewed by	Manfred Paier	AIT	26.08.2020
Approved by	Valeria Jana Schwanitz	HVL	02.12.2020
For delivery	Valeria Jana Schwanitz (revision of funding statement)	HVL	02.12.2020

Abstract/Publishable summary

This data management plan establishes guidelines to manage all data collected, created, processed and published in the H2020 project EERAdata. It also regulates the mechanisms to be used at the end of the project to share and preserve the project's data. To this end, the data management plan outlines all intended activities regarding data management and data governance. As a plan, this document is a living document, intended to be adapted and modified as required during the course of the project. As such, the data management plan will be hold in the project's wiki.

Suggested citation

Wierling, A., Schwanitz, V.J., Paier, M. (2020). Data Management Plan of the EU Project EERAdata, 2020 (Deliverable No. 30). Retrieved from the EERAdata website: https://eeradata.eu/about/deliverables.html

Disclaimer

This document contains a description of the EERAdata project findings, work, and products. Certain parts of it might be under partner intellectual property right (IPR) rules, so prior to using its content please contact the project principal investigator for approval: Valeria Jana Schwanitz, valerias@hvl.no.

In case you believe that this document harms in any way IPR, held by you or another person or representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent, and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is sole responsibility of the EERAdata consortium and can in no way be taken to reflect the views of the European Union.

Table of Content

Introduction

This data management plan establishes guidelines to manage all data collected, created, processed and published in the H2020 project EERAdata. It also regulates the mechanisms to be used at the end of the project to share and preserve the project's data. To this end, the data management plan outlines all intended activities regarding data management and data governance. As a plan, this document is a living document, intended to be adapted and modified as required during the course of the project. As such, the data management plan will be hold in the project's wiki (see below).

Section I - Administrative Information

This document is the data management plan for the H2020 project 'Towards a FAIR and open data ecosystem in the low carbon energy research community' (EERAdata), ID 883823. Parties responsible for ensuring the quality and fitness of data related to EERAdata are all partners in EERAdata. Specific responsibilities may be assigned in this document. All responsibilities described here are set in accordance with the regulations outlined in the Grant agreement (GA) and the consortium agreement (CA). In case of disagreement, GA and CA regulations take precedence. This data management plan complements and details other deliverables such as D1 (H-Requirement No.1), D2 (POPD – Requirement No.2) and D3 (NEC – Requirement No. 3).

EERAdata is taking part in the pilot on open research data (ORDP). This data management plan is one of the activities within this commitment. In addition, EERAdata will (a) deposit all non-administrative data in a research data repository, (b) ensure third parties can freely access, mine, exploit, and reproduce and disseminate your data. Information on how to implement these activities are given in the following sections. In particular, this document will provide related information and identify the tools needed to validate the results of the project.

EERAdata has assigned Dr. August Wierling, HVL, as data steward (DS) and data protection officer to coordinate all activities regarding the fitness and quality of all data, both as content as well as metadata.

Section II - Data and project management policies

Broadly, five different types of data are handled in the EERAdata project:

- Type A: Data collected, processed and created **during the series of workshops**, providing a comprehensive set of information regarding the FAIRification and opening of data leading towards a FAIR and open data ecosystem in the low carbon energy research community. This kind of data comprises workshop presentations, discussion notes, drafts for later publications, teaching material, guidelines, and short articles. All partners are involved in handling these data, mainly by their activities in WP2, WP4 – WP7. The responsibility for the governance and quality of the data align with the tasks in the work packages. Activities are coordinated in the context of WP2 and the overall coordination lies with HVL as WP2 leader.

- Type B: Data collected, processed, and **shared in the community platform for FAIR** and open energy data, which is designed and implemented in WP3. The community platform serves as an entry point for both, providers of energy data and users of energy data. It is conceived as a federated databases offering (a) advanced search functionalities and (b) enables harmonized access to external repositories. A community forum to discuss is connected to the community platform. Main responsibility lies with partner AIT as WP3 leader.

- Type C: A **pool of experts for FAIR data stewardship**. It is intended to register at least 60 domain experts in this activity. Being part of WP2, the responsibility for the pool of experts lies with HVL as WP2 leader.

- Type D: **Data connected to dissemination and communication activities** such as scientific publications, information on the website of the project (https://www.eeradata.eu/), and published in policy briefs and training material. All deliverables other than those already covered in types A-C are in this category. Being part of WP8, the responsibility lies with EERA as WP8 leader.

- Type E: **Project management data** such as minutes of board meetings, financial planning and reporting, internal discussions, etc. Here, the responsibility is with HVL as coordinating partner of the project.

All this data will be stored and processed in digital form. In order to accommodate to the different types of data, EERAdata has established different platforms to compile, store and publish the data. The choice has been made to ensure easy navigation, enabling versioning, and promoting shared authorship. In particular:

- A wiki application is holding all data of type A. The wiki can be accessed under https://eeradata-platform.eu/wiki/index.php?title=Main_Page. Using a wiki offers easy navigation of topics, linking of content, broad commenting and tracking of changes, collaborative authorship and a presentation layout known to many users. Page metadata will be added to ensure a FAIR use of all knowledge content created. As wiki software, the wide-used tool mediawiki is used offering a detailed access management. The wiki is hosted by a web hosting service provider.

- The community platform is implemented at the same web hosting service provider. More details will be provided in deliverable D10 'Functional specification'. A user support mechanism will be implemented on the community platform. Its actual form will be decided during the project, whereby two options are envisaged. The first option is to supply a technical guideline document for download. The second option offers dynamic overlays to give ad-hoc help (significantly more development effort). Both options would be supported by maintaining an email helpdesk. Moreover, a community forum on the same web hosting service will be embedded in the platform. Implementing an open source forum software (preferably phpBB, https://www.phpbb.com), user-account management will be supported similar to the wiki. It is at the core of the EERAdata project to offer detailed, community-agreed metadata information for all of the data served via the platform in agreement with the FAIR guiding principles. The specific form of this metadata is under development.

- Information on the pool of experts will be published on the wiki and the platform. Details on the interface will be decided during the course of the project. Further information on the type of information collected, consent forms etc. are available in deliverable D1.

- Dissemination and communication material will be published on the website eeradata.eu. All GDPR regulations will be implemented for accessing the web site.

- As a project management solution, the software *onlyoffice* is used. This European-based office software suite offers collaborative document editing, project management support, discussion forums, etc. Appropriate measures are implemented to guarantee access restrictions according to GDPR standards.

Section III - Creating, collecting and re-using data

Regulations on creating, collecting and re-using of data differ by the type of data involved:

- Data created of type A will be mainly original data of the EERAdata project. As part of the workshops, other databases will be analyzed and all licensing will be respected. A detailed documentation on these activities is part of the workshop documentation.

- Data created of type B are collected from external databases. All licensing will be respected. Data will be transferred from its original format into machine-readable JSON, which will then be handled by the TypeScript processing library. All agreements and regulations concerning licensing of databases sourced shall be added to this data management plan as an appendix.

- The pool of experts will be compiled during the course of the project. Related data will be made accessible via the community platform and the wiki. As described in D1, all GDPR requirements will be implemented.

- Documents and data created in the context of the management of the project (type E) will be created within the *onlyoffice* platform. A corresponding folder hierarchy and naming conventions for the files has been set up. Details can be found in the Project Management book.

Section IV - Curation of data

A regular backup of all data has a high priority. Both, the wiki and the project management solution offer easy possibilities for full as well as incremental backup. A full backup with storing data on the IT system of HVL is done once a month.

Section V - Publishing data and use of data during the project

In accordance with the commitment to the pilot on open research data, EERAdata will ensure that all research data that have been collected and produced can be ~~to~~ accessed, mined, exploited, reproduced and disseminated (free of charge for any user). To this end, reading access to the wiki is open to the public and all material can be downloaded.

If no rights of third parties are affected (as in the case of publishing in scientific journals), all material is published as creative commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0). For scientific publication, preference is given to publication as open access. As far as possible, all publication related data should be provided as supplementary material to the publication.

Intellectual property rights (IPR) of the platform are handled in the Project Management Handbook and the Consortium Agreement. However, EERAdata envisages that IPR about the concept for the platform are with the consortium, while IPR in regard to its implementation are with AIT. AIT makes the source code available, e.g., under the MIT license.

Section VI - Preserving data after the end of the project

During the project duration, the community platform will be hosted and maintained by AIT and finally migrated to ENEA's HPC facilities for future institutional hosting (installation on HPC platform, including data tools, to be discussed). By the end of the project, all data of type A-D will be integrated into the community platform as additional documentation. All contents of the website will be archived on the community platform.

Section VII - Giving access to data after the end of the project

The same policy as described under Section V also holds for access to the data of type A-D after the end of the project. Permanent access is ensured via the institutional hosting by partner ENEA. In addition, data of type A will be made accessible on the repository hvl.dataverse.no.