

HORIZON-CL5-2022-D3-01-11

Demonstration of innovative forms of storage and their successful operation and integration into innovative energy system and grid architectures



AGISTIN

Advanced Grid Interfaces for
innovative Storage INtegration

D1.2: Data management plan, V1

Document properties

| | |
|------------------------|---------------------------------------|
| FUNDING PROGRAM | HORIZON EUROPE |
| GRANT AGREEMENT NUMBER | 101096197 |
| PROJECT | AGISTIN |
| DELIVERABLE ID | D1.2 |
| TITLE | <i>D1.2: DATA MANAGEMENT PLAN, V1</i> |
| DISTRIBUTION LEVEL | <i>PU</i> |
| DUE DATE | 30.06.2023 |
| DATE SUBMITTED | 30.06.2023 |
| VERSION | V1 |
| WORK PACKAGE | WP1 |
| LEAD PARTICIPANT | EPRI |
| TOTAL NUMBER OF PAGES | 17 |

All intellectual property rights are owned by the AGISTIN consortium members and are protected by the applicable laws. Except where otherwise specified, all document contents are: “© AGISTIN project - All rights reserved”. Reproduction is not authorised without prior written agreement.

The commercial use of any information contained in this document may require a license from the owner of that information.

All AGISTIN consortium members are also committed to publish accurate and up to date information and take the greatest care to do so. However, the AGISTIN consortium members cannot accept liability for any inaccuracies or omissions, nor do they accept liability for any direct, indirect, special, consequential or other losses or damages of any kind arising out of the use of this information.

Views and opinion stated in this report reflects the opinion of the authors and not the opinion of the European Commission. The European Union is not liable for any use that may be made of the information contained in this document.

Version history

| VERSION | DATE | COMMENT |
|---------|------------|------------------------------|
| V1 | 30.06.2023 | FIRST VERSION OF AGISTIN DMP |
| | | |
| | | |



EXECUTIVE SUMMARY

This deliverable describes the identified data categories used within the AGISTIN project by the different project partners as well as ethical aspects and data handling requirements.

For the description of the data categories, each data set is broken down into the type of data and it is explained for what purpose the data is collected or generated. The source of the data is evaluated, and the confidentiality assigned. For cases of publishing project outcomes, for instance data sets, also the usefulness to external parties is explained and if already known repository named. At this stage of planning, however, not more details are available.

For ethical data handling and collection, relevant local, national, and international regulations, laws and agreements must be investigated and observed. This document is the first step to develop plans to obtain and document the informed consent of stakeholders regarding the use of their personal data. With these points in mind, the AGISTIN project is carefully considering how it will handle any personal data we might use during the project.

At the date of writing this report, all work packages are at an early phase in the planning of their data requirements or have not started their work yet. Detailed planning of the data use aspects will be undertaken during the first year of the project and the monitoring of the ethical use of the field trial data will continue during the project lifetime.

Therefore, an updated version of this deliverable will be released in project month 18.

Table of Contents

| | |
|---|----|
| Abbreviations and Acronyms | 1 |
| 1 Introduction..... | 2 |
| 1.1 Objectives of the work reported in this deliverable..... | 2 |
| 1.2 How to read this document | 2 |
| 2 Methodology..... | 3 |
| 2.1 FAIR Data | 3 |
| 2.1.1 Making data findable..... | 4 |
| 2.1.2 Making data openly accessible | 4 |
| 2.1.3 Making data interoperable | 4 |
| 2.1.4 Increase data re-use | 4 |
| 3 Data Summary..... | 5 |
| 3.1 Personal Data..... | 5 |
| 4 Allocation of Resources..... | 7 |
| 5 Data Security..... | 8 |
| 6 Ethical aspects and the AGISTIN policies related to data handling..... | 9 |
| 6.1 EU legislation related to our work..... | 9 |
| 6.2 Ethical Self-Assessment as made in the Description of Action..... | 9 |
| 6.2.1 Ethical dimension of the objectives, methodology and likely impact..... | 9 |
| 6.2.2 Compliance with ethical principles and relevant legislations..... | 9 |
| 6.3 Implementing the AGISTIN policy on personal data and privacy..... | 10 |
| 6.4 Handling of personal data for project mailing lists..... | 10 |
| 6.5 Handling of personal data collected as part of surveys conducted by AGISTIN | 11 |
| 7 Time plan for update | 12 |
| 8 References | 13 |

LIST OF TABLES

| | |
|---|----|
| Table I – Factsheet for Personal Data dataset..... | 5 |
| Table II - Ethical issues and relevance to the AGISTIN project..... | 10 |

Abbreviations and Acronyms

| Abbreviations and Acronyms | |
|----------------------------|--|
| CSV | Comma-Separated Values |
| DMP | Data management plan |
| DoA | Description of Action |
| EC | European Commission |
| EU | European Union |
| FAIR | Findable, accessible, interoperable, re-usable |
| GDPR | General Data Protection Regulation |
| ORDP | Open Research Data Pilot |
| R&D | Research and Development |

1 Introduction

Data management is an important task in every research project. Since Horizon 2020 and continuing with Horizon Europe programmes, European projects are required to adhere to the Open Research Data Pilot (ORDP). There are two main pillars to the Pilot: developing a Data Management Plan (DMP) and providing open access to research data, if possible. AGISTIN project will, as much as possible, comply with the ORDP requirements by creating and updating a project's Data Management Plan, initiated with this document, and by pledging to publish, when possible, scientific information according to the FAIR principle in publicly accessible research data repositories. The FAIR principle is explained in detail in section 2.1.

The AGISTIN project will identify best practices for a DMP following the work of previous H2020 and Horizon Europe projects. In the first revision of the Data Management Plan (DMP), we will present a clearly defined structured process for our DMP, taking as references the best practices identified from other H2020 and Horizon Europe projects and material provided by the European commission. These sources were used throughout the AGISTIN DMP:

- H2020 Programme, Guidelines on FAIR Data Management in Horizon 2020 (1)
- OpenAIRE H2020 project (2)
- Data management, H2020 Online Manual (3)

1.1 Objectives of the work reported in this deliverable

The AGISTIN DMP is submitted in month 6 of the project. The objective of this deliverable is to identify the datasets which are applied in the AGISTIN project. It describes how these datasets will be processed and shared to support the Horizon Europe Open Research Data Pilot during the project's development and after the project's conclusion.

1.2 How to read this document

This document can be read independently of other AGISTIN deliverables.

2 Methodology

The Data Management Plans (DMP) is a formal description of the procedures of data handling during and after a project. A DMP describes the data management life cycle for the data to be collected, processed, and/or generated. By providing an assessment of data used in a project and a structured approach for aspects as naming conventions, metadata, and versioning, the DMP should also support data quality, efficient processing and sharing of data and to ensure it is soundly managed.

As part of making research data findable, accessible, interoperable, and re-usable (FAIR), the 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;
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).

All data will be treated according to EU legislation, and respective national laws.

The AGISTIN DMP is submitted in month 6 of the project. The objective of this deliverable is to define the project's approach to the management of all the datasets which will be identified and collected in the scope of the AGISTIN project. It describes how these datasets will be processed and eventually shared to support the Horizon Europe Open Research Data Pilot during the project's development and after the project's conclusion.

2.1 FAIR Data

The FAIR data principle is required to be used in EU-Projects by the "Guidelines on FAIR Data Management in Horizon 2020" (EC, 2016). It should support the exchange of scientific data and lead to knowledge discovery and innovation. The FAIR data approach is described by the acronym:

- Findable data: Clear naming and versioning of (meta-) data, easy to find by both humans and computers
- Accessible data: It is clearly specified how the data is made available, including needed tools, protocols, authentication, and authorization
- Interoperable data: The published data uses standards and vocabularies that allow interoperability with applications and workflows for analysis, storage, and processing
- Re-usable data: The goal of the FAIR is reusability; therefore, it is clearly defined when and for which duration data is made available and under which licensing the data was published

The AGISTIN consortium will publish data whenever possible under the FAIR principle. The FAIR data management approach enables the exchange of scientific data which fosters knowledge discovery and innovation.

2.1.1 Making data findable

AGISTIN is committed to publish project results during the project period and plans to make them available after the end of the project when possible. Making data findable is realized by choosing the proper platform/repository for storing and publishing, as well as assigning the relevant metadata.

2.1.2 Making data openly accessible

AGISTIN will, when possible, publish processed data as well as project results. At this stage of the project, the available repository or repositories are not yet finally selected. Right now, both the numerous and emerging general-purpose data repositories, such as Zenodo, GitLab or OpenAIRE, or well-curated special-purpose repositories are considered.

2.1.3 Making data interoperable

Interoperability is a relevant aspect in the AGISTIN project and will be addressed through the scope of the project in different Work Packages. In general, interoperability of AGISTIN data will be realized through the use of popular and suitable data formats and written explanation.

2.1.4 Increase data re-use

Re-use of data can be encouraged through clear licensing, prompt dissemination, archiving, and quality assurance. At the time of writing, licensing schema for AGISTIN data has not yet been selected. Depending on the data, the licencing might differ per data set or publication.

3 Data Summary

The AGISTIN project strives for a consistent and transparent data management. This will support the project management as well as assist the cooperation between the project partners. With a forward-looking data management, the exploitation of the project is facilitated for the project partners as well as of all potential external beneficiaries like research institutes, universities, or R&D departments of companies.

Due to the heterogeneity of the data, which is to be collected and generated within the AGISTIN project, this chapter is structured by a differentiation of the data in data categories. This allows a detailed assessment of the data collection and generation as well of issues of data privacy and security for each dataset.

The following sections of this document will present the datasets or categories that have been identified at this stage of the project:

- Personal Data

As of now, the first data set only relates to personal data and is considered confidential and will not be published.

All partners of the project consortium support the policy of open-source approaches. If possible, work and results will be published in an open-source manner. Publishable datasets will be considered to be published on open data platforms in addition to the AGISTIN platform.

Document collections such as OpenAIRE and Zenodo will be considered for the publication and collection of AGISTIN reports, results, research papers, etc.

3.1 Personal Data

Table 1 – Factsheet for Personal Data dataset

| Factsheet for dataset 1 | |
|---|---|
| Dataset name | Personal data |
| Dataset description <i>(general description, type of data)</i> | <p>These data will be collected for the AGISTIN newsletter and for the execution of the AGISTIN survey on the identification of functional requirements for industrial users. It includes participant's:</p> <ul style="list-style-type: none"> • Name, surname, title • Contact information incl. company • Date and time of registration |
| Purpose of data collection/generation, relation to project objectives <i>(Why is the data needed?)</i> | <ul style="list-style-type: none"> • The personal data used for registering in the AGISTIN newsletter is needed to keep members updated on developments around the AGISTIN project and beyond. • EPRI collects personal data to distribute the newsletter and eventually to invite for events such as workshops or conferences and for continuously providing information related to the project for interested parties (separate consent will be asked). |
| Source of the data <i>(e.g., data from predecessor projects or from other</i> | <ul style="list-style-type: none"> • Newsletter members will be asked to provide the needed data when registering to the AGISTIN newsletter. The participants to the AGISTIN survey on the identification of functional requirements |

| | |
|--|--|
| <p><i>business activities, system or database from which the data is extracted)</i></p> | <p>for industrial users have been given the possibility to opt in for the newsletter subscription, by providing a specific voluntary consent during the survey execution.</p> <ul style="list-style-type: none"> • Consortium members, in particular those involved in dissemination and communication activities, will collect personal data at exhibitions, trade shows, and other relevant events. |
| <p>Format of the data <i>(e.g., database format like .csv, .xml)</i></p> | <ul style="list-style-type: none"> • List of newsletter subscribers will be securely stored in CSV format. |
| <p>Data security and privacy <i>(publishable or confidential?)</i></p> | <p>These data are strictly confidential and will be treated as personal data under European and national laws for personal data protection.</p> |
| <p>Data utility, usefulness to external parties <i>(To whom external might the data be useful e.g., for research?)</i></p> | <p>Personal data is confidential and will not be published.</p> |
| <p>Availability (long-term storage) <i>(If the data is published, where will it be published and for how long)</i></p> | <p>These data are strictly under personal data protection (European and national laws). The detailed description of the handling and protection of personal data is presented in Chapter 6 of this report.</p> |

4 Allocation of Resources

In order to make AGISTIN data FAIR, costs for storage, provision and data security as well as the creation of the necessary infrastructure to make it available are expected in the first place.

The costs for the development and maintenance of this environment were estimated during the application process and are included in the budget of the partner EPRI. As described before, parts of the data will be hosted in collaboration with other initiatives / organisations, therefore it is expected that not all datasets will be hosted directly on the AGISTIN resources.

In addition to the publication of the data on the AGISTIN resources, a publication on relevant open source and open data platforms is aimed for, the financing of which is not the responsibility of the AGISTIN project.

The partner responsible for the Data Management Plan is EPRI, supported by all other partners. Part of the work in AGISTIN is to develop a roadmap for the AGISTIN results to be sustainable and self-sufficient, even beyond the project's duration. A long-term preservation and availability of the datasets will be part of the to be developed roadmap. In addition, it is expected that with publishing data on open data platforms it will be available for a long period of time after the end of the project.

5 Data Security

Each partner is responsible for the security, recoverability, and storage of their own generated data (according to their institution or company practice).

In terms of hosting joint project data, it is stored on a Microsoft SharePoint installation made available by EPRI. This service is secured, controlled with username, password, domain filtering and specific roles for users within the project. Access is only granted by EPRI after a personal check by the Project Coordinator. The repository is accessible via a web application, which is SSL-secured with a Microsoft Corporation certificate, as well as the Microsoft Teams application. The data is hosted on a Microsoft server managed by the Microsoft Corporation located in Europe.

6 Ethical aspects and the AGISTIN policies related to data handling

The relevant EU legislation as well as the AGISTIN policy on personal data and privacy will be described in the following section. Relevant national legislations for each partner must be complied with by all project consortium members but will not be described in this deliverable.

6.1 EU legislation related to our work

The planned use of data conforms to current legislation and regulations in the countries where the project will be carried out. In particular, the project consortium members shall adhere to relevant EU legislation such as:

- The Charter of Fundamental Rights of the European Union;
- Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data;
- Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

6.2 Ethical Self-Assessment as made in the Description of Action

With this Data Management Plan being a public deliverable, but the Description of Action (DoA) of AGISTIN being confidential to project partners and commission, the ethical self-assessment section related to Personal Data from the DoA is included in this section.

6.2.1 Ethical dimension of the objectives, methodology and likely impact

Personal data will not be managed except for those needed for the direct implementation of the activities of the project. Data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation will not be collected neither managed.

Collected data will not have any impact neither for the individuals nor for their organisations. Data will be managed according to the data principles described in section 'Data Management and Management of other research outputs' of the AGISTIN proposal. Ethical and data protection concerns will be considered, fully complying with privacy issues.

6.2.2 Compliance with ethical principles and relevant legislations

The AGISTIN project will ensure research data follows the Findable, Accessible, Interoperable and Reusable data principle to make possible that knowledge is integrated and available for re-use in future projects. Data sharing will be addressed, taking into consideration any ethical and data protection concerns and fully complying with privacy issues.

6.3 Implementing the AGISTIN policy on personal data and privacy

The AGISTIN project fully supports and agrees to take the legal and ethical issues into consideration in order not to violate the privacy and individual rights of participants and stakeholders involved in the project.

The main legal restrictions concerning AGISTIN activities are regulated by countries legislation on data protection. Personal information which identifies an individual is strictly protected by the European Union legislation and must be accessed through specific permission. Table II below summarises our considerations related to the use of personal data in the AGISTIN project.

Table II - Ethical issues and relevance to the AGISTIN project

| Ethical issue | AGISTIN relevance | Ethical consideration |
|---|--|---|
| Does your research involve personal data collection and/or processing? | Yes, from AGISTIN subscribing to newsletter and participating to public events or activities, like workshops, conferences, surveys, etc. | <p>Informed consent is ensured if personal data is to be collected or processed.</p> <p>Identification information concerning individuals is kept within the organisation collecting it and is not provided to any organisation external to the AGISTIN consortium or to other AGISTIN partners.</p> <p>Reports and publications shall generalise findings and any personal data described in reports will be anonymised.</p> |
| Does it involve tracking or observation of participants? | No, neither tracking nor observation of participants is planned in the project. | <p>Should any need to tracking or observation of participants arise, their informed consent will be ensured.</p> <p>Identification information concerning individuals are not given to any organisation external to the AGISTIN consortium. Reports and publications will generalise findings and any participant data described in reports will be anonymized.</p> |
| Does your research involve further processing of previously collected personal data? | No, no further processing of collected data apart from the primary intended use is planned. | Should any need arise, their informed consent will be ensured. |

6.4 Handling of personal data for project mailing lists

We will develop project mailing lists to publicise our project results and activities. We will ensure that personal data, such as email addresses, provided to us for our project mailing lists, will only be used for the purposes explicitly agreed to by the participants and that it will not be given to third parties for other uses.

The AGISTIN Data Protection Officer will ensure that appropriate written confirmation of informed consent is obtained and that ethical and legal recommendations for the different activities and AGISTIN processes involving the collection, storage, processing, sharing, distribution and

destruction of personal and company related data are implemented.

6.5 Handling of personal data collected as part of surveys conducted by AGISTIN

Surveys are planned to be conducted over the course of the AGISTIN project, though it is not planned to collect personal data. If in any case, personal data shall be collected, AGISTIN will apply the following policy principles:

- Participants will be informed beforehand about the project aims and methods by written information and consent will be requested,
- At all times, participants will retain the right to withdraw consent and cease their involvement in the project without negative effects,
- Data coming from surveys will be treated anonymously in all cases where possible, and aggregated when needed to keep users' privacy, and
- Researchers will be informed that participating users have the right to remain anonymous. Where a questionnaire is sent through the post to subjects, their return of the questionnaire may be taken to imply consent.

7 Time plan for update

The first version of the AGISTIN DMP will be submitted in month 6 of the project time. The DMP is, however, a living document and will be subject to revisions and updates through the whole duration of the project. Updates can be done at need, when for example new data sets are identified in the project, or in a regular basis. The first planned update of the DMP will be in month 18 of the project, at the end of the first reporting period. One additional update will be planned for month 30 of the project, at the end of the second reporting period. The last update will be before the final project review in month 48, included in the final report.

8 References

1. EUROPEAN COMMISSION, Directorate-General for Research & Innovation. Guidelines on FAIR Data Management in Horizon 2020. [Online] 26 July 2016. [Riportato: 30 June 2023.] https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
2. OpenAIRE - Homepage. *OpenAIRE*. [Online] [Riportato: 30 June 2023.] <https://www.openaire.eu/>.
3. European Commission;. Data management, H2020 Online Manual. *Open Access Data Management*. [Online] [Riportato: 30 June 2023.] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm.