D1.2 Data Management Plan

October 2024 (M6)

Co-funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under the grant agreement № 101160145.

Deliverable Title	D1.2 Data Management Plan
Project acronym and Nº	SynGRID - Grant Agreement Nº: 101160145
Contractual Delivery Date	October 2024 (M6 of the project)
Actual Delivery Date	October 2024 (M6 of the project)
Work package	WP1
Responsible	Jan Jeriha
Lead beneficiary	IRI
Authors	Jan Jeriha, Tamara Smolej, Edin Lakić
Dissemination level	DMP-PU
Abstract	This DMP adheres to the FAIR principles, ensuring that project data is Findable, Accessible, Interoperable, and Reusable. Data will be made "as open as possible, as closed as necessary," balancing accessibility with ethical, legal, and security requirements. Datasets will be deposited in Zenodo for long-term preservation and public access, where appropriate. Strict security measures will be applied to protect sensitive information and comply with privacy regulations. Internally, all project data will be stored and shared via a secure SharePoint repository to ensure controlled access and collaboration among project partners. Ethics: Ethical considerations, including data privacy and consent, are integral to the project. Data collection and processing will strictly follow GDPR and other relevant regulations, ensuring that the rights and privacy of individuals are protected. Future updates to the DMP will further refine data handling and ethical compliance as the project progresses.
Keywords	Data Management Plan (DMP), Findable, Accessible, Interoperable, and Reusable (FAIR), Public, Zenodo

Disclaimer

The work described in this document has been conducted within the SynGRID project. This document reflects only the SynGRID Consortium view, and the European Union is not responsible for any use that may be made of the information it contains.

Copyright statement

Control Versions

Version	Date	Description	Author
V0.0	01.08.2024	New document	Jan Jeriha (IRI)
V0.1	27.09.2024	Complete document ready for Internal review	Jan Jeriha (IRI)
V0.2	04.10.2024	Internal Review performed	Edin Lakić, Tamara Smolej, Jan Jeriha (IRI)
V0.3	21.10.2024	First final version of the document ready	Edin Lakić, Tamara Smolej, Jan Jeriha (IRI)
V0.4	29.10.2024	Review performed by FER and integrated into the final version	Tomislav Antić (FER), Hrvoje Pandžić (FER), Jan Jeriha (IRI)

Executive Summary

The Data Management Plan (DMP) is a dynamic document that outlines the data management strategies adopted by the SynGRID consortium for research data throughout the project. As the SynGRID project progresses, the DMP will be regularly updated to reflect advancements, ensuring that the guidelines and recommendations align with the principles of making data Findable, Accessible, Interoperable, and Reusable (FAIR). This ongoing adaptation will support knowledge exchange, facilitate innovation, and maximize the impact of the research.

This first iteration of the DMP outlines the strategic approach to managing the data that will be generated or collected during the SynGRID project. It covers the types of data involved, the standards for data management, the preservation methods, and the accessibility of datasets for verification and reuse. These foundational efforts aim to ensure that the research outcomes contribute effectively to the innovation landscape, particularly in the context of low-voltage grid management and the integration of renewable energy sources (RES).

The structure and content of the DMP follow the European Commission (EC) guidelines as outlined in the document Guidelines on FAIR Data Management in Horizon 2020 [1]. Additionally, the recommendations offered by OpenAIRE [2] which instruct researchers to complete the DMP in full accordance with all regulations, were followed.

Table of Contents

Executive Summary 4
1 Introduction 7
1.1 Purpose of the document 7
1.2 Scope of the document 7
1.3 Structure of the document 7
2 Data Overview 9
2.1 Objectives of Data Management and connection to the project 9
2.2 Types, Standards and Formats of Dataset 9
2.3 Data Reuse 9
2.4 Predicted Timetable for DMP Updates 10
3 FAIR Data 11
3.1 Ensuring Data Findability with Metadata Support 11
3.2 Ensuring Data Accessibility 12
3.2.1 Publicly Available Repository of Research Data 12
3.2.2 Internal Repository of SynGRID 12
3.3 Open Access Publications 13
3.4 Making Data Interoperable 14
3.5 Improve Data Repurposing 14
3.6 Ensuring Deliverables Accessibility 15
4 Resource Allocation Strategy 16
4.1 Responsibilities and Decision Making 16
5 Data Protection and Confidentiality 18
5.1 Public Data Storage - Zenodo 18
5.2 Non-Public Data Storage - SharePoint 18
5.3 Data Privacy and Encryption 18
5.4 Data Protection Guidelines 18
6 Addressing Ethical Challenges 18
6.1 Procedures to protect data privacy 19
7 Conclusion 20
8 References and acronyms 21
8.1 References 21
8.2 Acronyms 21
9 Annex Data Management Plan 23

List of Figures

Figure 1: SharePoint software 13

List of Tables

Table 1: Timetable for DMP updates 10
Table 2: DC Levels of Interoperability [11] 23
Table 3: Overview of elements of DC Metadata Element Set [8] 23
Table 4: Best practices listed by European Data Portal [12] 23

1 Introduction

1.1 Purpose of the document

The purpose of this Data Management Plan (DMP) is to outline the approach that the SynGRID project will take to manage data through its lifecycle. This includes how data will be collected, processed, stored, and shared. It identifies key actions and strategies to ensure that research data are of a high-quality, secure, sustainable, and – to the extent possible – accessible and reusable. The DMP ensures that all data generated within the project follows the FAIR principles, supporting the project's goals of enhancing regional innovation and fostering cross-border collaboration. The DMP not only outlines data handling procedures but also ensures compliance and collaboration among all consortium members.

1.2 Scope of the document

The current version of this document serves as the first iteration, providing an initial framework for handling research data. It describes the types of data expected to be generated, the standards that will be applied, how data will be stored, and the portions of datasets to be shared for review or reuse. The DMP covers key aspects such as data findability, accessibility, interoperability, and reusability, while also addressing privacy, data security, and access via standard Application Programming Interfaces (APIs) and protocols. This document will guide the consortium's data management processes throughout the project, adapting as necessary to address emerging challenges, technological developments, and regulatory changes. While this initial version, submitted by month 6 of the project, may not provide detailed information on the specific datasets to be collected, generated, and processed, it aims to establish the foundation for an effective data management strategy throughout the research data lifecycle.

Future versions will expand on the topics outlined above, providing detailed descriptions of specific databases, usage conditions, and data management techniques. This ensures that data management methods are in line with SynGRID's developing aims, standards, and any new legal or regulatory obligations.

1.3 Structure of the document

The document is structured to guide the reader through the different aspects of data management within the SynGRID project. It outlines strategies for ensuring data is accessible, interoperable, and secure, while balancing openness with confidentiality. This document is organized into several key sections including data protection, ethical considerations, and the timeline for data release to both the consortium and the public:

- **Introduction**: Provides an overview of the document's purpose, scope, and structure (section 1).
- **Data Overview**: Discusses the guiding principles and the general approach to data management within the project (section 2).
- **FAIR Data**: It covers how the project will ensure that the data will be processed with FAIR principles, including metadata standards, licensing and data exchange protocols (section 3).
- **Resource Allocation Strategy:** Outlines the procedures for managing timelines for data release, and the methods for making data available to both the consortium and the public (section 4).
- **Data Protection and Confidentiality**: Addresses the measures taken to ensure data privacy, and the protection of sensitive information (section 5).

- **Addressing Ethical Challenges:** Examines the ethical considerations related to data management, including issues of consent, anonymization, and the handling of personal or sensitive data (section 6).
- **Conclusion**: Summarizes the key points and reiterates the commitment to balancing openness with the protection of commercial and confidential data (section 7).

2 Data Overview

2.1 Objectives of Data Management and connection to the project

The SynGRID project aims to enhance institutional and regional innovation in Widening countries by capitalizing on outcomes from Horizon 2020 and Horizon Europe projects. The data management objectives align with the project's broader goals of fostering collaboration, ensuring transparency, and maximizing the impact of research outcomes. The DMP is integral to the project as it supports these objectives by providing a structured approach to data handling, which is crucial for maintaining the quality and accessibility of the project's outputs. Proper and organized data management will facilitate quicker data exchange between partners, thereby enhancing the efficiency of project work.

2.2 Types, Standards and Formats of Dataset

The project will generate and handle various types of data, including experimental results, observations, and metadata from previous Horizon projects like COMPILE [3], X-FLEX [4], STREAM [5], and OPENTUNITY [6]. These datasets will be documented and stored using standard formats such as CSV (Comma Separated Value) for raw data, PDF for reports, and other relevant formats like JSON (Java Script Object Notation) or XML for metadata. Standards for metadata will follow the Dublin Core Metadata Initiative (DCMI) to ensure interoperability and ease of data integration across different platforms. Research data will be deposited in a trusted repository that follows relevant standards, and ensures long-term access.

The SynGRID project will contain several types of data, which include:

- **Research Data**: The project will gather a variety of research data, particularly on Low Voltage (LV) grid observability data and on the integration of renewable energy sources (RES) into low-voltage grids.
- **Validated Research Outcomes**: The project will also utilize and build upon validated research data from previous Horizon 2020 and Horizon Europe projects, such as COMPILE, X-FLEX, STREAM, and OPENTUNITY. These datasets have been pre-validated and are related to grid management, the integration of RES, and technological innovations in grid observability and controllability.
- **Technology and Environmental Data**: This includes data from technology environment analysis and regulatory gap analysis, focusing on the needs and developments within Distribution System Operators (DSOs) in the regions of Croatia, Greece, and Slovenia. Data related to the operational requirements of DSOs, such as those concerning grid planning and load management, will also be included.
- **Project and Collaboration Data**: Data generated from collaboration activities, including cross-border research networks, policy recommendations, and capacity-building activities, will also be collected. This data supports the objectives of fostering regional innovation and developing sustainable energy policies.

2.3 Data Reuse

Data reuse involves utilizing the data collected not only for the project's immediate objectives but also for future research, policy-making, and educational purposes. The data, particularly from previous projects and grid management activities, will be made available for other research initiatives, potentially leading to new innovations in renewable energy integration and grid management. This ensures that the valuable insights gained during the project are not confined to SynGRID alone but can contribute to broader scientific and technological advancements.

The core focus of SynGRID is the exploitation of results from other successful EU and national projects. Since we will not be generating new data, our aim is to replicate existing research data and leverage it for new project proposals. Data from existing projects will be managed securely, with access restricted to only what is necessary and only when required. Reused data, will be anonymized to safeguard personal information, ensuring that all reuse complies with ethical guidelines and legal requirements.

2.4 Predicted Timetable for DMP Updates

The DMP is structured to remain up-to-date and reflect the evolving needs of the project. While the initial DMP provided at the project's outset cannot establish a detailed list of all datasets, it will attempt to provide guidance for the following documents.
Subsequent updates are planned at key milestones, to incorporate new data types, adjust for changes in data handling practices, and integrate feedback from project partners. This approach ensures that the DMP supports the effective management and sharing of data throughout the project's lifecycle.
The DMP is a living document that will be updated according to the following timetable:

Table 1: Timetable for DMP updates

DMP Version	Delivery month	Description
D1.2	M6	Data Management Plan 1
D1.3	M18	Data Management Plan 2
D1.6	M36	Data Management Plan 3

3 FAIR Data

Effective data management is essential to maximizing the impact and value of research. To promote transparency, reuse, and collaboration, the FAIR data principles provide a framework for managing and sharing data in a way that enhances its long-term accessibility and utility.

The definition describes key components of the term FAIR data as [1]:

- **Findability**: Ensuring data discoverability through unique identifiers and trusted repositories,
- **Accessibility**: Managing IPRs based on data type, setting timelines for confidential data release, and explaining decisions to withhold data,
- **Interoperability**: Establishing common data, metadata and publication standards among project members,
- **Reuse of data**: Extensively exploring licensing options, with a preference for Creative Commons, to streamline and encourage data reuse.

With the implementation of the FAIR Principles in DMP, it is ensured that data of the project can be easily **accessed**, **understood** (implementation of common structure and usage of formal terminologies is required), **exchanged** and **reused**.

3.1 Ensuring Data Findability with Metadata Support

In the SynGRID project, identification and localization methods will be applied to the data being processed. The generated data will be identifiable and locatable through unique identification mechanisms. Files will be distinctly recognizable using standardized naming conventions and clear versioning. These naming conventions for documents and datasets were outlined in “D1.1 Project Management Handbook”.

It’s important to note that SynGRID will instead focus on replicating existing H2020 and Horizon Europe (HE) results and not focus on research and development but could still **generate original research data**. However, **proper cataloging and annotation of these replicated datasets with metadata**, as advocated by the European Data Portal [7], is still crucial. Metadata provides essential details about a dataset, such as its creation date, title, content, author, type, and size, making the data both human-readable and machine-readable. The European Data Portal strongly recommends using the **Data Category Vocabulary (DCAT)** Application Profile for metadata. Additionally, for datasets published on the internet, **Dublin Core (DC)** standards are applied. These standards not only help in making the data discoverable but also **improve clarity regarding its relevance, quality, and restrictions**. Publishing metadata as Linked Data further enhances discoverability, enabling better connectivity between users, software, and data across the internet. This structured metadata remains critical even when working with replicated datasets, ensuring they **are effectively integrated and accessible** within digital ecosystems.

The DC standards [8] were first standardized in 1998, even though invention of DC took place in 1995. They are represented by a set of 15 basic elements characterized by 4 levels of interoperability for data description. The levels are listed in Table 2 and elements are listed in Table 3 and are attached in the Annex Data Management Plan.
The elements presented in Annex Data Management Plan in section 9 are part of a larger set of metadata vocabularies and technical specifications. Due to the constant evolution of Data interoperability the DC standards are updated on a regular basis. In addition, all the 15 elements have been formally endorsed in the following standards:

- **ISO Standard 15836-2:2019 [ISO15836]** updating the original version from February 2009 [ISO15836],
- **ANSI/NISO Standard Z39.85-2012** of February 2013 [NISOZ3985],
- **IETF RFC 5013** of August 2007 [RFC5013].

The DCAT Application Profile, though based on (DC) standards, is not a vocabulary but rather a straightforward **specification for metadata descriptions** of EU governmental data and portals. Metadata will be published alongside the data in a **machine-readable format**, using standardized terms to define it. Key characteristics of the dataset, such as **local parameters, license, origin, and quality**, will be included in these descriptions. The Table 4 added to the Annex outlines the best practices recommended by the European Data Portal, which the SynGRID partners will adhere to when utilizing DCAT-AP.

3.2 Ensuring Data Accessibility

SynGRID is committed to the principle of "as open as possible, as closed as necessary," aiming to make research data openly accessible unless restricted by confidentiality, security, or ethical concerns. To achieve this, we will first identify all types of research data generated during and after the project. Based on this, a data accessibility analysis will follow, assessing any necessary software tools for accessing the data. While we intend to make data available wherever possible—provided privacy, ethical, or commercial concerns are not compromised—certain sensitive data may remain confidential due to security considerations. This analysis will be elaborated in future iterations of the DMP, and access to data will be granted to those with legitimate research interests.

As the project progresses, the following elements will be further specified for each dataset on a case-by-case basis, prior to making decisions regarding data management:

- The nature and scale of the data,
- The target audience and the level of interest,
- Availability of similar datasets and potential synergies,
- The potential for data integration and reuse by external researchers or users, and
- Any other relevant factors.

3.2.1 Publicly Available Repository of Research Data

To implement open access to research datasets, SynGRID will utilize Zenodo, a well-established European Open Research Repository Platform [9] that is fully integrated with OpenAIRE. The decision to upload and provide open access to a dataset in Zenodo will be made on a case-by-case basis, involving the PC (IRI UL) and the data-owning partner(s).

3.2.2 Internal Repository of SynGRID

Throughout the SynGRID project, **SharePoint** software will be used as the primary platform for managing project-related documentation. SharePoint will serve as a **centralized database for storing all reports, data, and other relevant documents**. As a project monitoring tool, it offers several valuable features that **enhance project tracking, transparency, and organization**. It is important to note that access to the internal repository is secured through user authentication, with permissions managed to ensure that only authorized personnel can access sensitive information.

Key features of SharePoint, used in SyngRID, include:

- **File Storage:**
  - Organizing files based on the Work Package (WP) and task structure.

- Implementing file version control to track updates and progress.

- **Team and Partner Management:**
  - Associating users with their respective partner organizations.
  - Maintaining contact details for all personnel involved.
  - Storing organizational information for each project partner.
- **Calendar:**
  - Providing an overview of key project activities, events, and deadlines.
- **Task Management:**
  - Creating task lists with detailed descriptions.
  - Assigning tasks to team members.
  - Defining start and due dates.
  - Setting reminders for task deadlines.
- **Milestones:**
  - Documenting detailed milestone descriptions.
  - Linking tasks to relevant milestones.
  - Assigning milestones to team members.
  - Using reminders and a milestone calendar to monitor progress.

In Figure 1, SharePoint software is presented.

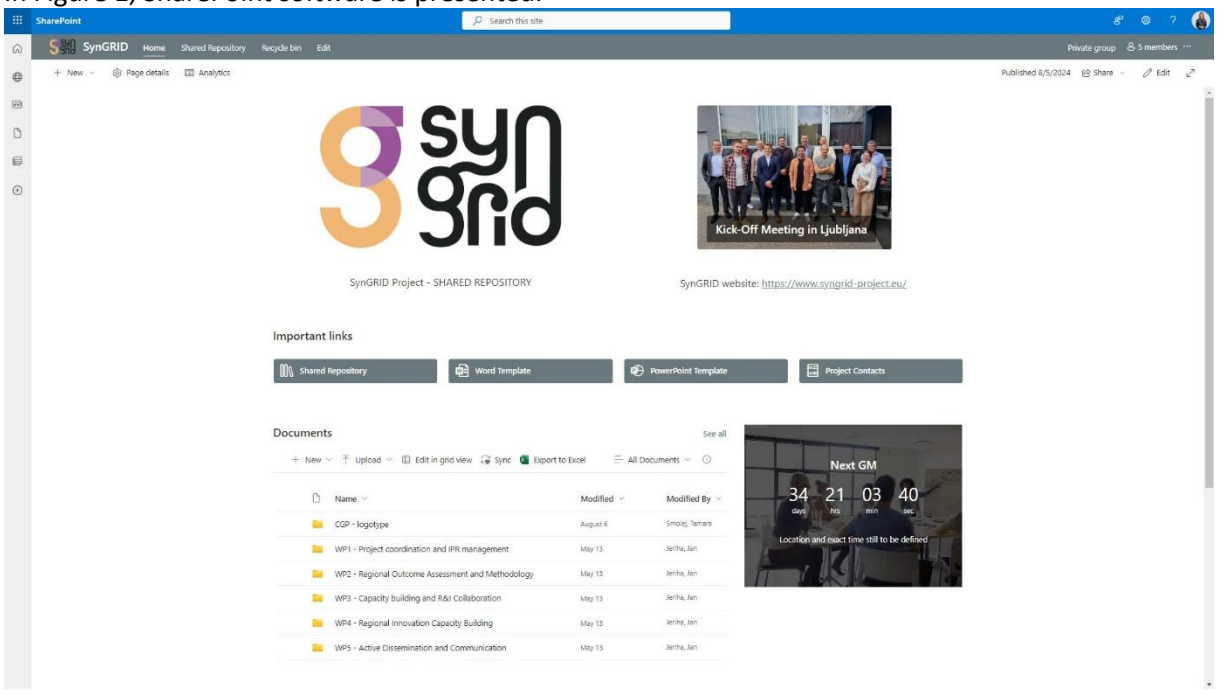

Figure 1: SharePoint software

3.3 Open Access Publications

The rules and principles of the EC Horizon 2020 [1] Framework Programme mandate that scientific results produced within H2020 projects must be made **openly accessible, ensuring broad dissemination of research findings**. This is in accordance with the guidelines outlined in the SynGRID Grant Agreement (section 1.2.4), ensuring **open access to all peer-reviewed scientific publications** related to SynGRID and its composite solutions. Additionally, the project consortium may decide to provide other outputs and results through open access channels.

According to Guidelines [10], there are two primary routes for providing open access to scientific peer-reviewed publications:

- **Self-Archiving** ('Green' Open Access): This method involves the researcher archiving the published article or final peer-reviewed manuscript in an online repository, either before, after, or alongside its formal publication. Access to the article may be **delayed by an 'embargo period' (typically six months)** as some publishers seek to recoup costs by selling subscriptions or charging for article access during this exclusivity window.
- **Open Access Publishing** ('Gold' Open Access): In this model, the article is **immediately made available in open access** by the publisher. The costs are shifted away from the readers and instead borne by the research institution to which the researcher is affiliated, or by the funding agency supporting the research.

These two methods **are not mutually exclusive** in EU-funded projects, meaning each SynGRID project beneficiary can choose the most appropriate approach for each publication. This will be further detailed in an updated version of the DMP.

At present, the primary strategy within the SynGRID project is parallel publishing. This approach **allows consortium members to publish in scientific journals** of their choice for maximum impact while ensuring open access to the results. Most academic journals support either **gold, green, or hybrid open access models**, giving SynGRID beneficiaries the flexibility to choose the most suitable publication venue. All publications will be assigned a DOI (Digital Object Identifier) to ensure they are easily discoverable and citable.

It's important to note that the open access requirement does not impose **an obligation to publish results**. As the guidelines clarify: "The decision to publish is **entirely up to the grant beneficiaries**. Open access becomes relevant only if publication is chosen as the method for dissemination." The decision-making process related to research results, and the potential paths for publication or other dissemination options, is illustrated in the accompanying graph.

Whenever possible, our research publications will be made available in open-access journals and shared on Zenodo. In exceptional cases, if conference publication necessitates a journal that is not publicly accessible, articles may be placed on personal or institutional websites. We will carefully consider access restrictions resulting from ethical or data protection concerns before publication and sensitive datasets will be pseudonymized.

3.4 Making Data Interoperable

An evaluation of data interoperability will be conducted, outlining which data and metadata vocabularies, standards, or methodologies will be adopted to ensure interoperability. This assessment will determine whether a standardized vocabulary will be applied across all data types in the dataset to support interdisciplinary interoperability. The starting point for this task involves **defining all types of data** to be managed during and after the project, as well as **identifying the components or actors** that will interact within the scope of the SynGRID project. As previously mentioned, no research and development will take place in SynGRID and therefore, there will be fewer interfaces required for data exchange.

3.5 Improve Data Repurposing

Data will be handled on a case-by-case basis throughout the project. Once a dataset is designated as public and made available on Zenodo, it will become fully reusable. It is also possible to apply an embargo period or restrict access to a specific whitelist of individuals, as outlined in Zenodo's policies.

Data will be accessible within and outside the consortium via standard APIs and protocols like REST. A publicly accessible repository adhering to **OpenAIRE principles and providing DOIs will be used for collaboration and citation**. While we aim to maximize data access, certain data may require anonymization or obfuscation before publication, including personal data, confidential information, and non-disclosure agreements among consortium members. These cases will be thoroughly documented.

Public datasets will be licensed under **Creative Commons Zero** (CC0) or **Open Data Commons Public Domain Dedication and License** (PDDL), enabling unrestricted reuse. Zenodo **guarantees the long-term preservation** of the final data. Once deposited in Zenodo, the data will be retained for the **lifespan of the repository**, which aligns with the lifetime of its host institution, CERN. Currently, CERN's experimental program is planned to continue for at least the next 20 years. All publicly accessible uploads on Zenodo will be securely stored in the same cloud infrastructure used **for research data from CERN's Large Hadron Collider**, utilizing CERN's robust repository software, INVENIO. This software also powers major repositories like INSPIRE HEP and the CERN Document Server.

The data will remain reusable for a **minimum of 20 years**, or until Zenodo ceases to host the datasets. The project also plans to implement a "data pedigree" approach, ensuring that all relevant information can **be traced back to the original data sources**. This data lineage, together with metadata, enables quality audits and sensitivity analysis of the project's outcomes.

3.6 Ensuring Deliverables Accessibility

After submission and approval by the EC, all the public project deliverables will be **made available through an organised and secure platform**, following established guidelines for openness and usability. Deliverables will be formatted according to widely accepted standards to ensure compatibility with **various tools and systems**, making them easily accessible to relevant stakeholders. Where possible, public deliverables will be **openly shared, while those containing sensitive or confidential information will be restricted to authorized users**, with clear access protocols in place. Regular reviews will be conducted to ensure compliance with accessibility requirements, facilitating seamless access to all necessary outputs.

4 Resource Allocation Strategy

The Resource Allocation Strategy in the SynGRID project is designed to ensure that all project tasks are effectively managed and completed within the set timeline and budget. It involves assigning specific roles and responsibilities to project partners, streamlining administrative and financial management, and addressing ethical considerations, IPR, and data management. Regular monitoring, communication among consortium partners, and close collaboration with stakeholders are key components to minimize risks and adapt to evolving project needs. The SynGRID project utilizes open-source software and free resources where possible, minimizing financial costs. However, expenses related to the hardware infrastructure, including servers to support the operation of SynGRID's technical frameworks and platforms, will be incurred. Additionally, working time for project setup, maintenance, and tool evolution, measured in person-months, will constitute a significant allocation of resources.

4.1 Responsibilities and Decision Making

As this DMP document is only an initial version, ongoing discussions will continue within the SynGRID consortium during the project to identify key deliverables and to determine appropriate methods for ensuring open accessibility where appropriate. To support this process, a dedicated time slot will be reserved at each project plenary meeting for updates and discussions related to data management.
The EC and project reviewers will be kept informed of the progress and related work on data management and open access. Regular updates will be provided through the project management reports, including details on any relevant publications and the implementation of the DMP. Individual responsibilities on data management in the project consortium are:

- **Project Coordinator (PC) (IRI):** IRI will lead data management, ensuring DMP compliance across all work packages. As PC, IRI will oversee data documentation, storage, and sharing, adhering to FAIR principles. They will manage the central repository, ensure backups and security, coordinate data management discussions during meetings, review data for open access, and report on DMP compliance to the EC and reviewers. The PC will be in control of all three versions of the DMP, which will be updated as agreed at major milestones.
- **Research roles (IRI, ICCS, FER, FERIT):** Research partners will be responsible for generating, validating, and securely managing datasets, ensuring compliance with FAIR principles and creating appropriate metadata. They will collaborate with the PC to share data through the central repository and open-access platforms while maintaining data security and confidentiality.
- **Technical, pilot and capacity building roles (FER, FERIT, PETROL):** These partners will manage data related to pilot site activities and capacity-building efforts. They will ensure data is properly collected, validated, and documented, particularly focusing on regional datasets. This partner will collaborate with other partners to ensure datasets are aligned with the DMP and are prepared for sharing through open-access platforms when applicable, while also ensuring data security and compliance with FAIR principles.
- **Dissemination Manager (DM) role (IRI):** The DM in the SynGRID project will be responsible for identifying project datasets and publications suitable for open access. They will ensure that data intended for public sharing is properly reviewed and prepared, coordinating with the PC and other partners. Additionally, the DM will manage the submission of datasets to open-access repositories and ensure that project outputs are widely accessible to external stakeholders, following the project's DMP and FAIR principles.

- **All partners:** Responsible for identifying and managing their project results and datasets suitable for publication. Ensure the quality, accuracy, and security of datasets before sharing them with the consortium or making them publicly accessible. Collaborate with the PC and DM on data sharing and publication processes.

5 Data Protection and Confidentiality

The SynGRID project prioritizes data safety and confidentiality by following high security requirements and regulations, such as the GDPR and applicable national guidelines. This section details the project's secure data storage, access control, encryption, and data sharing procedures. To ensure the secure storage and management of both public and non-public data within the SynGRID project, two primary repositories will be utilized: Zenodo and SharePoint.

5.1 Public Data Storage - Zenodo

For the storage of public data, Zenodo will provide robust security guarantees. Data files are backed up nightly, and file integrity is ensured by regular checksum verifications using the MD5 algorithm. In case of repository closure, Zenodo will strive to integrate all stored content into an alternative, suitable repository, ensuring data preservation.

5.2 Non-Public Data Storage - SharePoint

Non-public or sensitive project data will be hosted on a private, internal SharePoint server managed by PC at the University of Ljubljana (UL). Local backups of the data will be managed securely, with access restricted to authorized system administrators only. Files containing confidential information will be encrypted by the file owners using encryption tools (e.g., password-protected archives) before being uploaded to SharePoint. Secure web interaction will be ensured through HTTPS, and a Secure File Transfer Protocol (SFTP) will be made available as needed for data exchanges.

5.3 Data Privacy and Encryption

SynGRID will implement advanced encryption mechanisms to ensure robust data security, including server-side encryption for all stored data. This approach will provide a strong layer of protection, safeguarding sensitive information from unauthorized access or breaches. Access to such sensitive data will be strictly limited to designated project members, ensuring that even in the event of a security breach, unauthorized individuals will be unable to access the information. Additionally, where applicable, anonymization methods will be employed to further protect sensitive data, minimizing the risk of personal identification while maintaining the integrity of the research.

5.4 Data Protection Guidelines

To improve the security of project data, SynGRID will adopt several key guidelines. Whenever possible, anonymized and aggregated data will be used, minimizing reliance on individual-level data unless absolutely necessary. Sensitive data will be encrypted prior to any transfer outside the institution or project environment, ensuring it remains protected during transit. Data will be securely stored in at least two separate locations, reducing the risk of accidental loss. Additionally, the use of portable storage devices, such as USB flash drives, for transferring sensitive information will be minimized to limit potential security vulnerabilities. Consistent labelling of files will be enforced throughout the project to ensure uniformity and traceability at all stages of the data lifecycle.

6 Addressing Ethical Challenges

Through WP1, specifically Task 1.3 on "Ethics monitoring, IPR strategies and Responsible Research innovation", the project will establish a **robust framework for identifying, protecting, and managing intellectual property** (IP) generated during the project to ensure a **unified approach to IP management**.

6.1 Procedures to protect data privacy

Any data collection involving humans will be strictly held confidential at any time of the project. This means in detail that:

- All participants, including legal guardians when applicable, will be fully informed and allowed to provide consent for any monitoring and data collection activities. **Participation will be entirely voluntary, and detailed oral information will be provided to all participants**.
- **No** data will be collected **without the explicit informed consent of participants**. This ensures transparency about the research and guarantees that participants fully agree to the procedures by giving their informed consent.
- Personal or sensitive data **will not be stored centrally**. Additionally, where feasible, the data will be anonymized or abstracted in such a way that it will not impact the project's outcomes.
- Collected personal data will not be **sold or used for any purposes outside of the project**.
- A strict **data minimization policy will be enforced** throughout the project, overseen by an Ethics Panel. This ensures that no unnecessary data is collected.
- **Any unintended personal data collected during the research (shadow data) will be immediately deleted**. Efforts will be made to minimize the collection of such data. Special care will be taken to comply with the Council of Europe's Recommendation R (87), Article 2, which prohibits data collection based solely on racial origin, religious beliefs, sexual behaviour, political opinions, or membership in non-proscribed organizations, unless essential for a particular inquiry.
- Specific measures will be implemented **to protect participants from breaches of privacy and confidentiality, as well as from potential discrimination**. Participants' identities will remain confidential, and their involvement in the study will not be disclosed. Any incidental findings will be kept confidential and, upon request, erased from records.

The research will strictly follow the principles outlined in the EC *Ethics for Researchers: Facilitating Research Excellence* (2007). As the project involves the collection of user-related data, the consortium will fully comply with **all relevant European and national legislation**. This includes adherence to the **GDPR (EU) 2016/679**, **Directive 95/46/EC**, and **Directive 2002/58/EC**, which govern privacy, data protection, and the free movement of data. Additionally, compliance with the Universal Declaration of Human Rights and Convention 108 on the Protection of Individuals regarding the Automatic Processing of Personal Data, as well as any applicable national laws concerning sensitive information and data protection, will be ensured.

Data processing throughout the project will be conducted in strict accordance with **Article 7 of Directive 95/46/EC and GDPR (EU) 2016/679**. Personal data will only be processed under the following conditions: when the individual has provided informed consent, which will be obtained through detailed information sheets explaining the study's purpose, procedures, risks, and benefits; when processing is necessary to fulfil contractual obligations; to comply with legal requirements; or to protect the vital interests of the data subject. Consent will be documented via standardized forms, which will be securely stored in line with data protection regulations.

To **safeguard privacy**, all personal data collected during the project will be **anonymized, ensuring that individuals cannot be identified**. Data will only be processed for the specific purposes for which it was originally collected, with strong safeguards in place to protect personal information. Particular attention will be given to securing data, especially when stored for extended periods for historical, statistical, or scientific purposes, preventing unauthorized access, theft, or loss. A robust data management framework will be implemented to ensure compliance and the security of all personal data.

7 Conclusion

The document outlines the strategic approach to handling research data throughout the project lifecycle. By adhering to the FAIR principles, the project ensures that data is well-documented, discoverable, and usable by both current and future researchers. The implementation of Zenodo as the primary data repository reinforces the commitment to open science by providing a secure and sustainable platform for long-term preservation and dissemination of data. This open-access approach aligns with the principle of making data "as open as possible, as closed as necessary," ensuring a balance between transparency, reusability, and compliance with ethical and legal constraints.

Data protection and confidentiality will be maintained through compliance with legal frameworks, including GDPR, and by implementing security measures to ensure that sensitive data remains secure and accessible only to authorized personnel. Ethical considerations are paramount in managing sensitive data, and this plan prioritizes the protection of privacy, confidentiality, and intellectual property rights. Proper anonymization techniques and ethical guidelines will be followed to safeguard the rights and welfare of participants and collaborators.

It is important to highlight that this DMP is a living document. It will be updated and revised throughout the project's duration to reflect any changes in data management practices, technologies, or policies. This ensures that data management remains aligned with project goals and best practices in the evolving landscape of research data governance.

8 References and acronyms

8.1 References

[1] E. COMMISSION, "H2020 Programme Guidelines on FAIR Data Management in Horizon 2020," 2016. [Online]. Available: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

[2] OpenAIRE, "OpenAIRE Guides for Researchers How to comply with Horizon Europe mandate for Research DM," [Online]. Available: https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm.

[3] "Compile project," [Online]. Available: https://www.compile-project.eu/.

[4] "X-FLEX project," [Online]. Available: https://xflexproject.eu/.

[5] "STREAM project," [Online]. Available: https://stream-he-project.eu/.

[6] "OPENTUNITY," [Online]. Available: https://opentunityproject.eu/.

[7] European Data Portal, "Open Data Goldbook for Data Managers and Data Holders," January 2018. [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://data.europa.eu/sites/default/files/european_data_portal_-_open_data_goldbook.pdf.

[8] DCMI Usage Board, "Dublin Core™ Metadata Element Set, Version 1.1: Reference Description," 14 06 2012. [Online]. Available: https://www.dublincore.org/specifications/dublin-core/dces/.

[9] Zenodo, "General Policies v1.0," 2024. [Online]. Available: https://about.zenodo.org/policies/.

[10] ERC, "Guidelines on the Implementation of Open Access to Scientific Publications and Research Data in Projects supported by the European Research Council under Horizon 2020," November 2016. [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://erc.europa.eu/sites/default/files/ERC_Guidelines_Implementation_Open_Access.pdf.

[11] M. Nilsson, "Dublin Core," 01 05 2009. [Online]. Available: https://www.dublincore.org/specifications/dublin-core/interoperability-levels/. [Accessed 14 10 2024].

[12] Compile, "Deliverable 1.7: Data management plan (I)," 2018.

8.2 Acronyms

Acronyms list	
API	Application Programming Interfaces
CC0	Creative Commons Zero
CSV	Comma Separated Value
D	Deliverable
DC	Dublin Core
DCAT	Data Category Vocabulary

DCAT-AP	Data Category Vocabulary-Application Portal
DCMI	Dublin Core Metadata Initiative
DM	Dissemination Manager
DMP	Data Management Plan
DOI	Digital Object Identifier
DSO	Distribution System Operator
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
GDPR	General Data Protection Regulation
HE	Horizon Europe
HTTPS	Hypertext Transfer Protocol Secure
IP	Intellectual Property
IPR	Intellectual Property Rights
JSON	Java Script Object Notation
LV	Low Voltage
PDDL	Public Domain Dedication and License
PDF	Portable Document Format
PC	Project Coordinator
RES	Renewable Energy Sources
SFTP	Secure File Transfer Protocol
UL	University of Ljubljana
WP	Work Package
XML	Extensible Markup Language

9 Annex Data Management Plan

Table 2: DC Levels of Interoperability [11]

Dublin Core Levels of Interoperability	
Level 1: Shared term definitions	Shared vocabularies defined in natural language
Level 2: Formal semantic interoperability	Shared vocabularies based on formal semantics
Level 3: Description set syntactic interoperability	Shared formal vocabularies in exchangeable records
Level 4: Description set profile interoperability	Shared formal vocabularies and constraints in records

Table 3: Overview of elements of DC Metadata Element Set [8]

Contributor	An entity responsible for making contributions to the resource
Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant
Creator	An entity primarily responsible for making the resource
Date	A point or period associated with an event in the lifecycle of the resource
Description	An account of the resource
Format	The file format, physical medium, or dimensions of the resource
Identifier	An unambiguous reference to the resource within a given context
Language	A language of the resource
Publisher	An entity responsible for making the resource available
Relation	A related resource
Rights	Information about rights held in and over the resource
Source	A related resource from which the described resource is derived
Subject	The topic of the resource
Title	A name given to the resource
Type	The nature or genre of the resource

Table 4: Best practices listed by European Data Portal [12]

Publish the metadata together with the data
Use machine-readable formats to provide metadata
Use standard terms to define metadata
Describe the overall features of a dataset

Provide information about local parameters (date, time, value types, language)
Provide license information and information about its origin
Provide information about the quality