Solar Polygeneration Collector for Combined Heat, Power, Hydrogen Fuel and Wastewater Treatment

SPECTRUM

Data Management Plan

Prepared by:

Laboratório Nacional de Energia e Geologia, I.P.

March, 2025

Funded by the European Union

SPECTRUM has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101172891.

Disclaimer

Copyright Notice

Funded by
the European Union

Data Management Plan

Project Acronym	SPECTRUM
Project Title	Solar Polygeneration Collector for Combined Heat, Power, Hydrogen Fuel and Wastewater Treatment
Grant Agreement no.	101172891
Deliverable no.	D7.3
Dissemination Level	PU
Deliverable type	Report
Delivery date	March 2025
Lead beneficiary	LNEG
Author(s)	Cardoso, João P. (LNEG)
Reviewer(s)	Shalaby, Maryam (MG)
Suggested citation	Cardoso, J.P. Data Management Plan. Deliverable 7.3, SPECTRUM Project Consortium, 2025

Document history:

Version	Date	Status
V0.1	17/02/2025	Initial draft
V0.8	12/03/2025	Draft for review
V0.9	19/03/2025	Final draft after implementing comments, ready for editing
V0.9.1	21/03/2025	Final draft after editing, ready for approval
V1	27/03/2025	Final report

Funded by the European Union

Table of contents

Funded by
the European Union

List of tables

Funded by
the European Union

Nomenclature

Acronym	Meaning
CAD	Computer Aided Design
D	Deliverable
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Protection Regulation
JSON	JavaScript Object Notation
LNEG	Laboratório Nacional de Energia e Geologia, I.P.
WP	Work Package

Funded by
the European Union

Public Executive Summary

This document proposes a general approach to the handling of data resulting from the SPECTRUM project, containing guidelines and relevant information regarding the identification of the types of data that will be collected, processed and/or generated, the methodology and standards that will be applied for data organization, curation and preservation, and whether data will be shared/made open access and how it will be curated and preserved during and after the project conclusion.

Data from a wide variety of sources and with diverse types and formats are expected to be used and generated in the SPECTRUM project, including data from external sources (e.g.: scientific publications and databases, material properties databases, materials and component suppliers, etc.) and data generated in the internal activities of the projects (e.g.: laboratory testing, modelling and simulation activities, prototype construction and testing, communication and dissemination actions, etc.).

SPECTRUM will communicate and disseminate project results "as open as possible, as closed as necessary", adopting a data management methodology based on the FAIR (findable, accessible, interoperable and reusable) principle while maximizing the potential for exploitation of results by the project partners.

Data will be stored in two distinct repositories:

- An internal repository, created using the Microsoft SharePoint / Teams platform, is used for the collection, sharing and storage of all project data intended for internal use by the project partners;
- An external repository, Zenodo (https://zenodo.org/), is used for collecting, storing and publicly sharing data generated by the project classified as non-restricted.

Public deliverables and open-access publications will be made available on the project website and the Zenodo repository. Provisions for data security are in place according to the data location since the data generated by the project will be stored in different locations: at the partner level, consortium level, and public level.

The resources required for data generation, collection, and storage are foreseen in the existing Grant Agreement.

Funded by
the European Union

1. Introduction

The SPECTRUM project is a 42 months long research and innovation action funded by the European Union under the Horizon Europe programme, that aims to create innovative solar collectors for industrial applications by conceiving, developing, prototyping and testing lab-scale prototypes of solar collectors that fully harness the solar spectrum by converting solar radiation into three renewable energy vectors (solar heat, solar electricity and green hydrogen) while treating industrial wastewater.

As a result of the work carried out in Work Package (WP) 7 - Project coordination, management and quality assurance, this report, Deliverable 7.3 (D7.3), provides the first version of the Data Management Plan (DMP). This document proposes a general approach to the handling of data within the SPECTRUM project, during and after the end of the project, containing guidelines and relevant information regarding the identification of the types of data that will be collected, processed and/or generated, the methodology and standards that will be applied for data organization, curation and preservation, and whether data will be shared/made open access and how it will be curated and preserved during and after the project conclusion.

The DMP will be continuously updated throughout the project lifetime with intermediate reports published in months 15 (D7.13) and 27 (D7.14) and a final update by the end of the project in month 42 (D7.15).

Section 2 summarises the data reused and generated by the project. Section 3 describes how the FAIR (Findable, Accessible, Interoperable and Reusable) principles are applied to the project data. Section 4 discusses the allocation of resources required for data generation, collection and storage activities. Section 5 reviews the data security provisions implemented in the project. Section 6 reviews the ethical aspects related to the handling of project data. Finally, section 7 concludes the report by recapitulating its contents.

Funded by
the European Union

2. Data summary

Data from a wide variety of sources and with diverse types and formats are expected to be used and generated in the SPECTRUM project, including data from external sources (e.g.: scientific publications and databases, material properties databases, materials and component suppliers, etc.) and data generated in the internal activities of the projects (e.g.: laboratory testing, modelling and simulation activities, prototype construction and testing, communication and dissemination actions, etc.). The data produced by the project will be treated according to the FAIR principles (see section 3) and the general rules regarding data protection, namely the Regulation (EU) 2016/679 (General Data Protection Regulation, GDPR). The following subsections summarize the main characteristics foreseen for the data used and generated by the project.

2.1. Re-use of existing data

The project will use existing data from external sources and internal data from the project partners, leveraging existing knowledge and information to develop and validate the SPECTRUM project's new hybrid collectors' concepts. Different data sources, such as scientific publications and databases, material properties databases, and material and component suppliers, will be used. In all cases, the reuse of existing data will be referenced, ensuring traceability of its source.

It is foreseen to re-use existing data to drive the development of several project activities, such as the identification of relevant industrial wastewater streams, the design of new photocatalysts, the design of the collector prototypes and their sub-components, the study of the integration of the collectors in industrial applications, the assessment of life cycle impacts and feasibility of the SPECTRUM collectors.

2.2. Data types and formats

Many data types and formats are expected to be used and generated in the SPECTRUM project, including numerical values, graphics and text, from external sources (re-use of data) or from experimental testing, computational simulations, surveys and communication and dissemination activities.

Table 1 presents a preliminary summary of the data types and formats predicted to be used and generated in the SPECTRUM project.

Funded by
the European Union

Table 1 - Types and format of data foreseen to be generated and used in the SPECTRUM project

WP	Type of data	Data format
1	Industrial wastewater sources, Results from material characterization analysis, Results from experimental testing	Lists and tables (e.g.: xlsx, docx, csv, txt), Numerical data (e.g.: csv, txt, xlsx, docx), Images and graphics (e.g., jpg, png), Reports (e.g.:docx, pdf)
2	CAD models, Specifications of components, Simulation models and results	Simulation models (e.g.: tracepro, mph, py, ees, m), CAD files (e.g.: stl, stp, dwg), Numerical data (e.g.: csv, txt, xlsx, docx), Images and graphics (e.g., jpg, png), Reports (e.g.:docx, pdf)
3	CAD models, Prototype requirements and design, Testing protocols, Results from experimental testing	Lists and tables (e.g.: xlsx, docx, csv, txt), CAD files (e.g.: stl, stp, dwg), Numerical data (e.g.: csv, txt, xlsx, docx), Images and graphics (e.g., jpg, png), Reports (e.g.:docx, pdf)
4	Simulation models and results, Hydrogen separation, purification and storage requirements	Simulation models (e.g.: py, ees, m), Numerical data (e.g.: csv, txt, xlsx, docx), Images and graphics (e.g., jpg, png), Reports (e.g.: docx, pdf)
5	Results from life cycle and feasibility assessment, Survey results, List of legislations and regulations	Numerical data (e.g.: csv, txt, xlsx, docx), Images and graphics (e.g., jpg, png), Reports (e.g.: docx, pdf)
6	Project news, Photos and videos, Audience engagement statistics from social media and website	Text and images (docx, pdf, jpeg, png, mpeg), Exports from social media and website analytics (e.g.: xlsx, pdf)
7	Project management data such as financial and technical reports	Reports and deliverables (e.g.: xlsx, docx, pdf)

2.3. Purpose of data

Data, both generated within the project or retrieved from external sources, will be used to drive the project activities, namely, to support and validate the:

- Identification of industrial wastewater sources with high potential for photocatalytic remediation with co-generation of hydrogen;
- Development of low-cost, sustainable photocatalysts, optimizing their recovery, regeneration and reuse;
- Optimization of photocatalytic process parameters and the identification of the photocatalyst most suitable for use in the SPECTRUM prototypes;
- Design and development of the SPECTRUM hybrid solar collectors' sub-components, namely the concentrator, spectral splitter, photocatalytic reactor, thermal absorber and PV module;

Funded by
the European Union

- Integration of all sub-systems in a single device, optimizing its overall efficiency;
- Development and construction of two prototypes using standardized procedures for the factory testing of their components;
- Systematic testing of the prototypes to assess their performance;
- Development of control algorithms for optimum operation of the SPECTRUM-based systems;
- Development of the conceptual design of the up-scale SPECTRUM collectors and analysis of their potential application in selected industrial processes;
- Identification of safety requirements for SPECTRUM-based systems related to hydrogen, separation, purification and storage;
- Evaluation of the collectors' impact on environmental sustainability and their long-term techno-economic feasibility;
- Identification of potential legislative and regulatory barriers;
- Analysis of the social impact and improvement of the social awareness and acceptability of the novel technologies;
- Development of actions to communicate and disseminate the project activities and results, raising awareness about the developed technologies and related benefits to diverse audiences such as research and academia, industry actors, and policymakers;
- Provision of technical and scientific proof of the viability, benefits and scalability of the SPECTRUM collectors as a new technology to target audience such as research & academia, industry actors, and policymakers;

Moreover, project-generated data will also be used to coordinate all administrative and financial procedures of the project and to monitor and assess the project's technological aspects, quality, risks and progress.

2.4. Size of data

The total size of the data generated by the project is currently unknown, depending on several variables that can only be determined with the execution of the project (e.g. number of experimental tests, simulation runs, etc.). Therefore, it will be assessed only at the end of the project.

2.5. Potential users of data

The data generated by the project will be helpful to the scientific community working in solar photocatalysis for water treatment and/or hydrogen generation, solar photovoltaics, and solar thermal energy, and others. Moreover, it will be of interest to

Funded by
the European Union

companies working in industrial wastewater treatment and renewable energy generation, namely renewable hydrogen, electricity or thermal energy vectors, and their integration and use in industrial processes.

There is significant potential for exploitation of project results; therefore, the partners will adopt a proactive policy that pursues all possible patents and licensing opportunities arising from the project activities. Consequently, several deliverables are classified as sensitive, and their related data will not be publicised. Nevertheless, public executive summaries will be published for all deliverables.

Funded by
the European Union

3. FAIR data

SPECTRUM aims to communicate and disseminate project results "as open as possible, as closed as necessary", adopting a data management methodology based on the FAIR (findable, accessible, interoperable and reusable) principle.

3.1. Making data findable, including provisions for metadata

The use of persistent data identifiers or metadata for data shared or internal use between SPECTRUM project partners is not foreseen. Data available for external use will have digital object identifiers provided by the external, certified repository (Zenodo, http://zenodo.org), where the data will be stored and made available for external users. Moreover, project data stored in the certified repository will also be referenced on the project website.

To improve its discoverability, metadata will be used to classify and describe all data collected and generated by the project to be made available for external use and deposited in the external repository. There are no metadata standards covering the different scientific fields and data types generated by the project; therefore, the following minimum information should be included in the metadata:

- Date of data generation;
- Date of deposit;
- Author(s) and, if possible, organization(s);
- Description of contents, including data type (raw or processed), origin of the data (experimental, simulation, etc.), format of the data and names/description of data fields;
- Project name, acronym and reference number;
- Horizon Europe grant funding acknowledgement;
- Licensing terms of the dataset;
- Persistent identifiers for the dataset.

Metadata of deposited data will be open under a Creative Common Public Domain Dedication (CC 0).

3.2. Making data accessible

An internal repository, created using the Microsoft SharePoint / Teams platform, is used to collect, share and store all project data intended for internal use by the project partners.

Funded by the European Union

Data generated by the project, classified as non-restricted, will be collected, stored and publicly shared in the Zenodo repository (https://zenodo.org/) under open licenses defined by Creative Commons. Public deliverables and open-access publications will also be available on the project website. It is not foreseen to have an expiry date on data made publicly available.

The consortium will identify the data to be available openly in a case-by-case evaluation through an internal consultation and approval process. Data can be classified as restricted if it includes sensitive information or personal data protected by the GDPR or scientific and technological data the publication of which would undermine its exploitation or make more difficult the intellectual property protection by the project partners.

Access to open data will not be controlled and will follow the rules established by the selected trusted repository. Access to data in the internal repository will be restricted to persons belonging to the project team through the use of the authentication system in place by the Microsoft SharePoint / Teams platform (username and password).

3.3. Making data interoperable

SPECTRUM will implement provisions to ensure interoperability at the level of the dataset metadata, facilitating the discovery of the dataset and parsing of the data to facilitate its exchange and reuse. The use of Zenodo, as a public data repository, ensures that metadata is available using a JavaScript Object Notation (JSON) format following Zenodo JSON schema, ensuring data can be integrated with other related datasets, simplifying its reuse. Moreover, this metadata can be exported using standard formats such as MARCXML and Dublin Core.

3.4. Increase data re-use

The project beneficiaries will own and exploit data generated in the project after its completion. Open access to the public data generated by the project, which will be entered into the Zenodo repository, will be governed by Creative Commons licenses. It will be available to re-use according to the repository platform's terms of use, i.e., it will be retained for a minimum of 20 years, the expected minimum lifetime of the Zenodo repository at the time of writing.

Although not planned, it is possible to implement embargo periods for the data at a project beneficiary request and following a decision of the project consortium to allow for suitable exploitation of the data, namely for applications for protection of intellectual property rights.

Funded by
the European Union

4. Allocation of resources

Data generation, collection and storage fall within the activities foreseen in the existing Grant Agreement and are covered by the resources made available to all partners according to said agreement.

All partners bear responsibility for the quality and management of the data resulting from the project, namely the partners executing the actions generating the data and the corresponding work package leaders.

LNEG will support costs associated with the internal repository as project coordinator. This repository will be maintained for at least 5 years after project completion.

No costs are expected concerning the public repository as the selected repository, Zenodo, does not bear costs associated with long-term storage (minimum of 20 years) and data preservation, being supported by its own funding sources.

Funded by
the European Union

5. Data security

Data generated by the project will be stored in different locations: at the partner, consortium, and public levels. Provisions for data security are in place according to the data locations.

Data generated by project partners in the execution of their tasks will firstly be stored individually by each partner, subject to the rules, regulations, and procedures in place at each institution concerning data storage and security.

At the consortium level, data will be stored in the internal repository, Microsoft SharePoint / Teams platform, being subject to the security provisions implemented by this platform; in particular, only authorized users, as defined by the project coordinator, can access the repository, being this access password protected. The platform has a built-in versioning tool that allows for data recovery.

At the public level, data security follows the procedures established by Zenodo, including restricted access to the physical premises of the data centres, the use of HTTP protocol for access to the repository, and the use of cryptographic password hashing algorithms for user password protection.

Funded by
the European Union

6. Ethical aspects

The project was subjected to an ethics review during the proposal stage and grant agreement preparation, and no ethics issues were identified.

The project partners will exchange basic personal contact information to allow the consortium to meet and exchange emails and relevant documentation. The SPECTRUM consortium will implement the best practices to ensure data protection; namely, all will respect the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 and limit the personal data collection to (i) informed individuals, (ii) authorised retrieval, (iii) confidential treatment, storage and use of such data.

Data collected in surveys will be subject to the same practices; moreover, data will be anonymised, and personal data will be removed.

Data collected from the experimental tests and measurements will not contain personal data. Nevertheless, it will be subject to the same practices. External dissemination of data (raw data or treated data) will be performed under the FAIR data principles, and personal information will be removed from all data for external dissemination.

Funded by
the European Union

7. Conclusions

This deliverable describes the data management procedures to be adopted by the SPECTRUM project, addressing the data types and associated formats, the purpose of the data, its expected size and potential users. Moreover, this document describes the use of FAIR principles in managing data generated by the project, the required allocation of resources for its management, data security provisions and associated ethical aspects.

An update to the DMP will be published in December 2025 (month 15) as D7.13 – Data management plan - M15 update.

Funded by
the European Union