Plan Overview

A Data Management Plan created using DMPonline

Title: Setting Up green eNergy Research In SErbia (SUNRISE)

Creator: Aleksandra Lekic

Principal Investigator: Predrag Stefanov

Data Manager: Vaibhav Nougain, Jelena Stojković

Project Administrator: Predrag Stefanov, Jelena Stojković

Contributor: Aditya Shekhar, Jose Maria Maza Ortega, Marene Larruscain, Pablo Eguia, Kristina Džodić

Affiliation: Delft University of Technology

Funder: European Commission

Template: Horizon 2020 DMP

ORCID iD: 0000-0003-4054-2202

Project abstract:

The EU aims to reduce greenhouse gas emissions and be climate-neutral by 2050. Serbia has also adopted the Clean Energy for all Europeans package and agreed to a decarbonization roadmap. The transition to net-zero greenhouse gas emissions includes the decarbonization of the power sector. However, up to date, the majority of the energy in Serbia is fossil-fueled, with very little or no progress towards energy decarbonization. Moreover, pollution in Serbia is reaching alarming numbers, placing it as one of the leading countries in Europe related to bad air quality and pollution. Nonetheless, research capacity in the power system decarbonization domain at UB, as well as other actors in Serbia and the region, are undeveloped and insufficient for competitive R & I. Therefore, SUNRISE aims to support the Power System Department (the University of Belgrade – School of Electrical Engineering) in improving excellence capacity in the field of Power system decarbonization. Research excellence will be promoted in three
dimensions by improving research infrastructure through the development of Real-Time Simulation Laboratory for power system simulation, upskilling research staff through training and mobility, the development of new approaches in exploratory research projects with partners. Furthermore, as network and collaboration between UB and other research institutions in the EU is modest or non-existent, as well as low linkages between UB and industry, the special focus in SUNRISE will be given to mobility and networking activities. One more SUNRISE objective is to improve the research management and administrative skills of the staff at UB, to enhance funding attraction and participation in international research projects. SUNRISE partners have a long experience and high scientific impact in the power system decarbonization field, and they can support UB to improve scientific performance and research profile, as well as participation success in HE.

ID: 116077

Start date: 01-01-2023

End date: 31-12-2025

Last modified: 21-10-2025

Grant number / URL: 101079200

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Setting Up green eNergy Research In SErbia (SUNRISE) - Initial DMP

1. Data summary

Provide a summary of the data addressing the following issues:

- **State the purpose of the data collection/generation**
- **Explain the relation to the objectives of the project**
- **Specify the types and formats of data generated/collected**
- **Specify if existing data is being re-used (if any)**
- **Specify the origin of the data**
- **State the expected size of the data (if known)**
- **Outline the data utility: to whom will it be useful**

2. FAIR data

2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

2.2 Making data openly accessible:

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

2.3 Making data interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

2.4 Increase data re-use (through clarifying licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

3. Allocation of resources

Explain the allocation of resources, addressing the following issues:

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long term preservation**

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

This project will not lead to the creation of any personal data and therefore this question does not apply.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

Question not answered.

Setting Up green eNergy Research In SErbia (SUNRISE) - Detailed DMP

1. Data summary

State the purpose of the data collection/generation

SUNRISE is funded through the EU HORIZON-WIDERA-2021-ACCESS-03-01 action, under No. 101079200. This project is devoted to the twinning of good research and management practices from well-established EU universities to the widening country university. This project gathers the following partners:

- University of Belgrade (ETF), School of Electrical Engineering (Elektrotehnički fakultet), Serbia - project coordinator;
- Technische Universiteit Delft (TUD), Faculty of Electrical Engineering, Mathematics and Computer Science, Netherlands;
- University of Seville (US), Spain;
- University of the Basque Country UPV/EHU, Faculty of Engineering in Bilbao, UPV/EHU, Spain;
- The Business Technology Incubator of Technical Faculties Belgrade (BITF), Serbia.

In the project SUNRISE the data will be collected for the following purposes:

- Management and administrative - WP1, WP7;
- Research Management Office - rules and regulations, training material - WP6;
- Planning of the laboratory - WP2;
- Course material - WP3;
- Training material - WP3, WP4;
- Gather research and educational results - WP5.

WP3 will generate data related to the training activities and working visits which will be focused on the use of the real-time simulation and hardware-in-the-loop (HIL) tests. These data, i.e. simulation and HIL test results, will be accessible for the consortium members. In ths manner, the data will be also available as input of WP4 devoted to enhancing the R&I capacity through exploratory research projects.

Explain the relation to the objectives of the project

This DMP tackles all objectives set by the SUNRISE project:

1. Advancing the excellence capacity at the ETF in the energy transition toward decarbonization.

The outcome of this objective is setting up the Real-Time Simulation Laboratory at ETF. TUD, US, and UPV/EHU will contribute in terms of the provision of training for the ETF staff for real-time simulation in WP3. Virtual training as well on-site post-training during the short time visits will be organized for knowledge transfer and the exchange of best practices.

For this objective, data is produced in the design of real-time simulation platforms in RTDS and OPAL-RT. The training videos will be recorded for the MOOC. Furthermore, onsite training material about the use of real-time software will be created in form of RSCAD/RTDS, MATLAB/OPAL-RT, presentations, and recordings.

In addtion to this, the Real-Time Simulation Laboratory at ETF will be equipped with a Voltage Source

Converter (VSC) designed and built by US. Therefore, data produced in this part is related to all the design concerns: electrical schemes, mechanical integration schemes, VSC simulation, unitary testing of the main components and integrated testing.

2. Increasing the reputation and research profile of the ETF and its staff.

Development and training related to the Real-Time Simulation laboratory specifically dedicated to power sector decarbonization increase the potential to be recognized for its excellence and top scientific research activities. This goal can be measured by the number of publications (18), the number of masters (20) and Ph.D. (6) students, the number of research projects applications (3), and the funding attracted (0.7M€). Raised reputation at the international level can be assessed through international collaboration agreements with businesses and other research institutions (10) and a number of international visiting researchers (5). Mobility of research staff will be especially in focus with short-term staff exchanges (21), expert visits (12), and short-term on-site and virtual training.

For this objective data will be produced in the form of journal/conference publications which will be open-source and available using arXiv, LinkedIn, and gold open access subscription. Furthermore, the mobility of the staff, both academic and non-academic, will be described in the publicly available report.

3. Improving the research management and administrative skills of the staff at the ETF.

The concrete deliverable of this objective is setting up a research management office with a professional research support staff.

For this objective, ETF will create all rules and regulations for the functioning of the research management office. The administrative and research staff from ETF will be trained during the short-term visits to other beneficiaries. The training material will be offered in the form of presentations and text documents.

4. Strengthening the network between ETF and international research institutions and stronger linkages with national and regional companies.

To alleviate these obstacles, the objective of SUNRISE is to enhance the network of ETF and international research institutions, as well as, national and regional companies. Strengthening networking activities will include conferences and congresses attendance (9). Social media and networks, including ResearchGate, Academia.edu, Twitter, Instagram, and Facebook will be used to promote the research profile of the ETF and reach a wider audience. This moreover includes an audience beyond the power systems field. Also, a detailed website related to the project will be established. Two workshops with the target audience will be organized to facilitate outreach activities with industry partners. Also, two one-week summer schools will be organized from the project knowledge.

Specify the types and formats of data generated/collected

Administrative and management data:

- The figures would be in standard formats (.pdf, .jpg, .eps, .png).
- For reports will be used .doc(x) and .tex files.
- For management purposes, .xlsx and .csv files will be used.
- The communication started with Trello, but later on was moved to email communication and MS Teams. The data will be saved in the 4TU.ResearchData after ending the project.

Research and training data:

- In the case of large data quantitative datasets, we will use .csv and .xlsx.
- RSCAD simulations would use .dfx, while other simulations would be in .slx format. CBuilder uses .cpp and .h files.
- OPAL-RT simulations will be designed using MATLAB file formats, and C++ file formats .cpp and

.h.
- Typhoon HIL-related files: .cus, .tse, and .runx.
- The research manuscripts would be prepared using .doc(x) or .tex format.
- All metadata will be stored in .txt format.
- Electrical and mechanical schemes of the VSC prototype will be provided using .pdf format.
- PCB details used in the VSC prototype will be provided using the corresponding Gerber files.

Specify if existing data is being re-used (if any)

For training and the course material will not be used already created data and therefore this question does not apply.

Specify the origin of the data

For training and the course material will not be used already created data and therefore this question does not apply. Therefore, all data will be created during the project following the project's distribution of IP rights.

State the expected size of the data (if known)

- 1 - 10 GB

Data generated by the project is expected to be in the GB range (up to ten GB). We do not expect to store data above 100 GB. Data created by the consortium will be stored and maintained by 4TU.ResearchData and GitHub.

At first, for project management, a non-professional version of Trello was used. Later, the project coordination was switched to MS Teams and email communication.

Outline the data utility: to whom will it be useful

The generated data will be useful for:

- Entities in the widening countries - showing good practices in the execution of the twinning project with all public deliverables.
- Research and development institutions, and industry - with all public conference and journal publications.
- Real-time laboratories - with the generated real-time simulation scenarios, and hardware-in-the-loop tests.
- Students and researchers - with created training material and recorded MOOC.

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

All data-supporting publications will be made openly available through 4TU.ResearchData. This is a trusted and certified research data repository (it has a Data Seal of Approval certification) to store and reuse applied technical-scientific data. In addition, to further aid their discoverability, keywords describing the datasets will be added. Data sets will have a DOI and will therefore always be findable and accessible.

Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

The project SUNRISE will use EngMeta (schema.org metadata), meaning that all datasets are indexed in Google Dataset Search. Every dataset would be assigned a Digital Object Identifier (DOI), to make them citable and persistently available.

Outline naming conventions used

Documents and archives will follow the naming convention:
First letters: SUNRISE
Underscore:
Next letters: Deliverable number [Dx.y] [x=WP number, y=deliverable number]
Underscore:
Next letters: Short explanatory title for the document
Underscore: _
Next letters: "v" and number of revisions of this specific report [v0.1=draft version, v1.0=final version]
The title will be followed with the suitable extension.

Deliverable document's full title will look as follows:
[SUNRISE_Dx.y_Title_v0.1] example: SUNRISE_D1.1_Project Handbook_v0.1.{extension}

Outline the approach towards search keyword

- The work project titles and relevant words will be used as keywords.
- All databases will have 'readme' files, where the description of the repository and its use will be provided.
- Databases will be covered with the installation tutorials and user manuals containing definitions of variables, units of measurement, any assumptions made, the format and file type of the data, and software used to collect and/or process the data.

Outline the approach for clear versioning

Since the design of the metadata archive is a repetitive process, project beneficiaries will clean unnecessary files in each iteration. GitHub is a known platform that allows the versioning of files.

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

We are not aware of any metadata standards in our area. For that matter, we will create 'readme' files which will contain information about each folder.

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

- Simulation test cases (.dft or .slx).
- Simulation outputs considering different scenarios (.xlsx or .csv).
- Codes (.mat, .py, .txt).
- Text and figures in reports, publications, and presentations (.docx, .pdf, .eps).

All codes, simulation test cases, presentations, and videos will be stored using 4TU.ResearchData and GitHub. They will be kept private until accepted for publication and stored for 15 years after finishing the relevant deliverable and task.

Specify how the data will be made available

- ETF - All the codes and data related to the simulations and HIL tests will be available on GitHub. The public data version will be included in the common project repository (4TU.ResearchData). All the research papers will be publicly available in the publisher's website (Gold open access).
- TUD - All simulation codes and output data will be made available on a dedicated GitHub version provided by TUD and will be made available after the related publications are accepted. The public version of the data will appear in 4TU.ResearchData repository.
  All published manuscripts will be freely accessible through the publisher's website (Gold open access) or TU Delft's repository (Green open access). Wherever appropriate for the journal and academic discipline, preprint servers (e.g. arxiv.org) will be used to disseminate working papers and submitted manuscripts at an early stage.
- US - All the codes and data related to the simulations and HIL tests will be available on GitHub. The public version of the data will be included in the common project repository (4TU.ResearchData). All the research papers will be publicly available on the publisher's website (Gold open access) or a US repository (Green open access).
- UPV/EHU - All the codes and data related to the simulations and HIL tests will be available on GitHub. The public version of the data will be included in the common project repository (4TU.ResearchData). All the research papers will be publicly available on the publisher's website (Gold open access) or UPV/EHU's additional repository (Green open access).

Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

- *Simulation data*: .xlsx and .csv files can be accessed using Microsoft Excel, Google Sheets and compatible software.
- *Text and figures:* any .docx or .pdf files can be accessed using Micrsoft Word, Google Docs, Adobe, Acrobat reader and compatible software.
- *Simulation cases*: .dft files can be accessed using RSCAD software, .slx simulation, and .mat code files can be accessed using MATLAB/SIMULINK (also used by OPAL-RT simulations).

All stable versions of the software code will be placed in GitHub repositories as publicly available and followed with the updated 'readme' file.

Specify where the data and associated metadata, documentation and code are deposited

The developed software cases and codes presented in academic papers will be shared on 4TU.ResearchData portal and via GitHub. This way, they will be publicly available to anyone for re-use under an open license. They will also be assigned a Digital Object Identifier (DOI) to make them citable and persistently available.

We will mention the version of the software release and the dependencies for the creation and use of the data.

Specify how access will be provided in case there are any restrictions

There will not be restricted data and thus, this question does not apply.

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

- ETF - The code produced by ETF, along with relevant documentation to facilitate its reusability, will be added to the shared project repository (4TU.ResearchData).
- TUD - The code produced by TUD will be placed in 4TU.ResearchData and explainable for later re-use. The training and course material will be publicly available and self-explanatory.
- US. The code produced by US will be placed in the common project repository (4TU.ResearchData) including appropriate documentation for facilitating its re-use. The training materials will available for the consortium members being self-explanatory.
- UPV/EHU. The code produced by UPV/EHU will be placed in the common project repository 4TU.ResearchData, including appropriate documentation for facilitating its re-use. The training materials will available for the consortium members being self-explanatory.

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

Project SUNRISE will use standard vocabulary for all data types present in the data set. This will allow inter-disciplinary interoperability.

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

- Project deliverables will be completely public.
- Stable versions of the simulation data will be saved using a CC BY 4.0 license that allows:
  - **Share** — copy and redistribute the material in any medium or format;
  - **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.
  - All datasets will be accompanied by rich and descriptive metadata, including a DOI, compliant with DataCite metadata schema, to ensure that all datasets are findable and accessible online.
- For the code, we will use the MIT license.
- Journal and conference publications will be offered as open-access or as open-access preprints.

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

Data will be available for re-use after finishing the task and deliverable for which they are created. Furthermore, certain data will be kept private until being accepted for publication for a maximum of 12 months after finishing the relevant deliverable and task.

Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

There is no restriction for data re-use as we will not collect user data, while simulations scenarios, code, and output will be based on open-source data or data available in the public domain.

Describe data quality assurance processes

To ensure the good quality of data, all produced and collected data will be checked by the task lead, work package lead and the project coordinator.

Specify the length of time for which the data will remain re-usable

Data will remain reusable for at least 15 years after finishing the project, which is guaranteed by 4TU.ResearchData. The data on GitHub will remain public even after 15 years.

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

Project beneficiaries have allocated funds for this task:

- ETF has allocated funds for publications in an open-access format. No funds will be required for archiving since TU Delft is providing the project repository free of charge.
- TUD has allocated funds for MOOC - Organisation of the recording of the MOOC, but also received a donation from IEEE IES. 4TU.ResearchData is available for TUD; therefore, no funds will be needed for archiving. TUD has allocated funds for open source publishing.
- The US has allocated funds for publications in an open-access format. No funds will be required for archiving since TU Delft is providing the project repository free of charge.
- UPV/EHU has allocated funds for open-access publications. No funds are required for archiving since TU Delft provides the project repository free of charge.

Clearly identify responsibilities for data management in your project

- The responsibility for data management relies on the project coordinator ETF.
- Data collection and sorting rely on the beneficiaries responsible for the task in which those data are generated. Furthermore, work package leaders are responsible that all data gathered in the task is well grouped inside the whole work package.

Describe costs and potential value of long term preservation

Project administration and management:

- Trello free version - no costs;
- Site maintenance - subcontracted by ETF: costs 1000 €;
- LinkedIn page - without costs;
- Facebook page - without costs;
- Twitter page - without costs.

Open-access publishing:

- arXiv - publishing without costs;
- MOOC - recording placed on TUD server, 4TU.ResearchData, without costs;

- journals - paid per publication, depending on the APC rate of the selected journal. TUD has green open access to several journals, and in cases when TUD staff are co-authors of the publication, these costs will not apply. The US and UPV/EHU have gold open access to some journals in the power system field. Therefore, the publication costs do not apply when the US or UPV/EHU researchers are co-authors of the publication. ETF has allocated funds for Gold Open access in journals.
- 4TU.DataResearch will be responsible for the long-term preservation of the data. TUD is offering access to all project partners.

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

We will not use sensitive data. We will use the default security measures defined by the project board. In this case, only team members can access the designated server. The storage security is ensured by the project coordinator.

Each beneficiary should ensure internal storage and appropriate security measures:

- ETF will use the institutional, personal accounts and store data in the ETF MS Teams.
- TUD will use Surfdrive for internal share, and for the storage of the consortium files.
- US researchers will use the institutional, personal accounts provided by Microsoft.
- UPV/EHU will use the UPV/EHU MS team's SharePoint folder and the internalGorde Talde tool

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

This project will not lead to the creation of any personal data and therefore this question does not apply.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

All consortium partners will store data following 4TU.ResearchData platform upon the completion of the project.

At UPV/EHU, data will be stored internally usingGorde Talde platform.

Setting Up green eNergy Research In SErbia (SUNRISE) - Final review DMP

1. Data summary

State the purpose of the data collection/generation

SUNRISE is funded through the EU HORIZON-WIDERA-2021-ACCESS-03-01 action, under No. 101079200. This project is devoted to the twinning of good research and management practices from well-established EU universities to the widening country university. This project gathers the following partners:

- University of Belgrade (ETF), School of Electrical Engineering (Elektrotehnički fakultet), Serbia - project coordinator;
- Technische Universiteit Delft (TUD), Faculty of Electrical Engineering, Mathematics and Computer Science, Netherlands;
- University of Seville (US), Spain;
- University of the Basque Country UPV/EHU, Faculty of Engineering in Bilbao, UPV/EHU, Spain;
- The Business Technology Incubator of Technical Faculties Belgrade (BITF), Serbia.

In the project SUNRISE the data will be collected for the following purposes:

- Management and administrative - WP1, WP7;
- Research Management Office - rules and regulations, training material - WP6;
- Planning of the laboratory - WP2;
- Course material - WP3;
- Training material - WP3, WP4;
- Gather research and educational results - WP5.

WP3 will generate data related to the training activities and working visits, which will be focused on the use of the real-time simulation and hardware-in-the-loop (HIL) tests. These data, i.e., simulation and HIL test results, will be accessible for the consortium members. In this manner, the data will also be available as input for WP4, devoted to enhancing the R&I capacity through exploratory research projects.

Explain the relation to the objectives of the project

This DMP tackles all objectives set by the SUNRISE project:

1. Advancing the excellence capacity at the ETF in the energy transition toward decarbonization.

The outcome of this objective is setting up the Real-Time Simulation Laboratory at ETF. TUD, US, and UPV/EHU will contribute in terms of the provision of training for the ETF staff for real-time simulation in WP3. Virtual training as well as on-site post-training during the short-term visits will be organized for knowledge transfer and the exchange of best practices.

For this objective, data is produced in the design of real-time simulation platforms in RTDS and OPAL-RT. The training videos will be recorded for the MOOC. Furthermore, onsite training material about the use of real-time software will be created in the form of RSCAD/RTDS, MATLAB/OPAL-RT, presentations, and recordings.

In addition to this, the Real-Time Simulation Laboratory at ETF will be equipped with a Voltage Source

Converter (VSC) designed and built by the US. Therefore, data produced in this part is related to all the design concerns: electrical schemes, mechanical integration schemes, VSC simulation, unitary testing of the main components, and integrated testing.

2. Increasing the reputation and research profile of the ETF and its staff.

Development and training related to the Real-Time Simulation laboratory, specifically dedicated to power sector decarbonization, increases the potential to be recognized for its excellence and top scientific research activities. This goal can be measured by the number of publications (18), the number of master's (20) and Ph.D. (6) students, the number of research project applications (3), and the funding attracted (0.7M€). Raised reputation at the international level can be assessed through international collaboration agreements with businesses and other research institutions (10) and a number of international visiting researchers (5). Mobility of research staff will be especially in focus with short-term staff exchanges (21), expert visits (12), and short-term on-site and virtual training.

For this objective, data will be produced in the form of journal/conference publications, which will be open-source and available using arXiv, LinkedIn, and a gold open-access subscription. Furthermore, the mobility of the staff, both academic and non-academic, will be described in the publicly available report.

3. Improving the research management and administrative skills of the staff at the ETF.

The concrete deliverable of this objective is setting up a research management office with a professional research support staff.

For this objective, ETF will create all rules and regulations for the functioning of the research management office. The administrative and research staff from ETF will be trained during the short-term visits to other beneficiaries. The training material will be offered in the form of presentations and text documents.

4. Strengthening the network between ETF and international research institutions and strengthening linkages with national and regional companies.

To alleviate these obstacles, the objective of SUNRISE is to enhance the network of ETF and international research institutions, as well as national and regional companies. Strengthening networking activities will include conference and congress attendance (9). Social media and networks, including ResearchGate, Academia.edu, Twitter, Instagram, and Facebook, will be used to promote the research profile of the ETF and reach a wider audience. This, moreover, includes an audience beyond the power systems field. Also, a detailed website related to the project will be established. Two workshops with the target audience will be organized to facilitate outreach activities with industry partners. Also, two one-week summer schools will be organized from the project knowledge.

Specify the types and formats of data generated/collected

Administrative and management data:

- The figures would be in standard formats (.pdf, .jpg, .eps, .png).
- For reports, .doc(x) and .tex files will be used.
- For management purposes, .xlsx and .csv files will be used.
- The communication started with Trello, but later on was moved to email communication and MS Teams. The data will be saved in the 4TU.ResearchData after ending the project.

Research and training data:

- In the case of large data quantitative datasets, we will use .csv and .xlsx.
- RSCAD simulations would use .dfx, while other simulations would be in .slx format. CBuilder uses .cpp and .h files.
- OPAL-RT simulations will be designed using MATLAB file formats, and C++ file formats .cpp and .h.

- Typhoon HIL-related files: .cus, .tse, and .runx.
- The research manuscripts would be prepared using .doc(x) or .tex format.
- All metadata will be stored in .txt format.
- Electrical and mechanical schemes of the VSC prototype will be provided using the .pdf format.
- PCB details used in the VSC prototype will be provided using the corresponding Gerber files.

Specify if existing data is being re-used (if any)

For training and the course material will not be used, already created data, and therefore, this question does not apply.

Specify the origin of the data

For training and the course material will not be used, already created data, and therefore, this question does not apply. Therefore, all data will be created during the project following the project's distribution of IP rights.

State the expected size of the data (if known)

- 1 - 10 GB

Data generated by the project is expected to be in the GB range (up to ten GB). We do not expect to store data above 100 GB. Data created by the consortium will be stored and maintained by 4TU.ResearchData and GitHub.

At first, for project management, a non-professional version of Trello was used. Later, the project coordination was switched to MS Teams and email communication.

Outline the data utility: to whom will it be useful

The generated data will be useful for:

- Entities in the widening countries - showing good practices in the execution of the twinning project with all public deliverables.
- Research and development institutions, and industry, with all public conference and journal publications.
- Real-time laboratories - with the generated real-time simulation scenarios, and hardware-in-the-loop tests.
- Students and researchers - with created training material and recorded MOOC.

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

All data-supporting publications will be made openly available through 4TU.ResearchData. This is a trusted and certified research data repository (it has a Data Seal of Approval certification) to store and reuse applied technical-scientific data. In addition, to further aid their discoverability, keywords describing the datasets will be added. Data sets will have a DOI and will therefore always be findable and accessible.

Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

The project SUNRISE will use EngMeta (schema.org metadata), meaning that all datasets are indexed in Google Dataset Search. Every dataset would be assigned a Digital Object Identifier (DOI) to make it citable and persistently available.

Outline naming conventions used

Documents and archives will follow the naming convention:
First letters: SUNRISE
Underscore:
Next letters: Deliverable number [Dx.y] [x=WP number, y=deliverable number]
Underscore:
Next letters: Short explanatory title for the document
Underscore: _
Next letters: "v" and number of revisions of this specific report [v0.1=draft version, v1.0=final version]
The title will be followed by the suitable extension.

Deliverable document's full title will look as follows:
[SUNRISE_Dx.y_Title_v0.1] example: SUNRISE_D1.1_Project Handbook_v0.1.{extension}

Outline the approach towards search keyword

- The work project titles and relevant words will be used as keywords.
- All databases will have 'readme' files, where the description of the repository and its use will be provided.
- Databases will be covered with the installation tutorials and user manuals containing definitions of variables, units of measurement, any assumptions made, the format and file type of the data, and the software used to collect and/or process the data.

Outline the approach for clear versioning

Since the design of the metadata archive is a repetitive process, project beneficiaries will clean unnecessary files in each iteration. GitHub is a known platform that allows the versioning of files.

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

We are not aware of any metadata standards in our area. For that matter, we will create 'readme' files which will contain information about each folder.

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

- Simulation test cases (.dft or .slx).
- Simulation outputs considering different scenarios (.xlsx or .csv).
- Codes (.mat, .py, .txt).
- Text and figures in reports, publications, and presentations (.docx, .pdf, .eps).

All codes, simulation test cases, presentations, and videos will be stored using 4TU.ResearchData and GitHub. They will be kept private until accepted for publication and stored for 15 years after finishing the relevant deliverable and task.

Specify how the data will be made available

- ETF - All the codes and data related to the simulations and HIL tests will be available on GitHub. The public data version will be included in the common project repository (4TU.ResearchData). All the research papers will be publicly available on the publisher's website (Gold open access).
- TUD - All simulation codes and output data will be made available on a dedicated GitHub version provided by TUD and will be made available after the related publications are accepted. The public version of the data will appear in 4TU.ResearchData repository.
  All published manuscripts will be freely accessible through the publisher's website (Gold open access) or TU Delft's repository (Green open access). Wherever appropriate for the journal and academic discipline, preprint servers (e.g., arxiv.org) will be used to disseminate working papers and submitted manuscripts at an early stage.
- US - All the codes and data related to the simulations and HIL tests will be available on GitHub. The public version of the data will be included in the common project repository (4TU.ResearchData). All the research papers will be publicly available on the publisher's website (Gold open access) or a US repository (Green open access).
- UPV/EHU - All the codes and data related to the simulations and HIL tests will be available on GitHub. The public version of the data will be included in the common project repository (4TU.ResearchData). All the research papers will be publicly available on the publisher's website (Gold open access) or UPV/EHU's additional repository (Green open access).

Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

- *Simulation data*: .xlsx and .csv files can be accessed using Microsoft Excel, Google Sheets, and compatible software.
- *Text and figures:* any .docx or .pdf files can be accessed using Microsoft Word, Google Docs, Adobe, Acrobat Reader, and compatible software.
- *Simulation cases*: .dft files can be accessed using RSCAD software, .slx simulation, and .mat code files can be accessed using MATLAB/SIMULINK (also used by OPAL-RT simulations).

All stable versions of the software code will be placed in GitHub repositories as publicly available and followed by the updated 'readme' file.

Specify where the data and associated metadata, documentation and code are deposited

The developed software cases and codes presented in academic papers will be shared on 4TU.ResearchData portal and via GitHub. This way, they will be publicly available to anyone for reuse under an open license. They will also be assigned a Digital Object Identifier (DOI) to make them citable and persistently available.

We will mention the version of the software release and the dependencies for the creation and use of the data.

Specify how access will be provided in case there are any restrictions

There will not be restricted data, and thus, this question does not apply.

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

- ETF - The code produced by ETF, along with relevant documentation to facilitate its reusability, will be added to the shared project repository (4TU.ResearchData).
- TUD - The code produced by TUD will be placed in 4TU.ResearchData and explainable for later re-use. The training and course material will be publicly available and self-explanatory.
- US. The code produced by US will be placed in the common project repository (4TU.ResearchData), including appropriate documentation for facilitating its re-use. The training materials will be available for the consortium members, being self-explanatory.
- UPV/EHU. The code produced by UPV/EHU will be placed in the common project repository 4TU.ResearchData, including appropriate documentation for facilitating its re-use. The training materials will be available for the consortium members, being self-explanatory.

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

Project SUNRISE will use standard vocabulary for all data types present in the data set. This will allow inter-disciplinary interoperability.

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

- Project deliverables will be completely public.
- Stable versions of the simulation data will be saved using a CC BY 4.0 license that allows:
  - **Share** — copy and redistribute the material in any medium or format;
  - **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.
  - All datasets will be accompanied by rich and descriptive metadata, including a DOI, compliant with the DataCite metadata schema, to ensure that all datasets are findable and accessible online.
- For the code, we will use the MIT license.
- Journal and conference publications will be offered as open-access or as open-access preprints.

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

Data will be available for re-use after finishing the task and deliverables for which they were created. Furthermore, certain data will be kept private until being accepted for publication for a maximum of 12 months after finishing the relevant deliverable and task.

Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

There is no restriction for data re-use, as we will not collect user data, while simulation scenarios, code, and output will be based on open-source data or data available in the public domain.

Describe data quality assurance processes

To ensure the good quality of data, all produced and collected data will be checked by the task lead, work package lead, and the project coordinator.

Specify the length of time for which the data will remain re-usable

Data will remain reusable for at least 15 years after finishing the project, which is guaranteed by 4TU.ResearchData. The data on GitHub will remain public even after 15 years.

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

Project beneficiaries have allocated funds for this task:

- ETF has allocated funds for publications in an open-access format. No funds will be required for archiving since TU Delft is providing the project repository free of charge.
- TUD has allocated funds for MOOC - Organisation of the recording of the MOOC, but also received a donation from IEEE IES. 4TU.ResearchData is available for TUD; therefore, no funds will be needed for archiving. TUD has allocated funds for open source publishing.
- The US has allocated funds for publications in an open-access format. No funds will be required for archiving since TU Delft is providing the project repository free of charge.
- UPV/EHU has allocated funds for open-access publications. No funds are required for archiving since TU Delft provides the project repository free of charge.

Clearly identify responsibilities for data management in your project

- The responsibility for data management lies with the project coordinator ETF.
- Data collection and sorting rely on the beneficiaries responsible for the task in which that data is generated. Furthermore, work package leaders are responsible for ensuring that all data gathered in the task is well grouped within the whole work package.

Describe costs and potential value of long term preservation

Project administration and management:

- Trello free version - no costs;
- Site maintenance - subcontracted by ETF: costs 1000 €;
- LinkedIn page - without costs;
- Facebook page - without costs;
- Twitter page - without costs.

Open-access publishing:

- arXiv - publishing without costs;
- MOOC - recording placed on TUD server, 4TU.ResearchData, without costs;
- journals - paid per publication, depending on the APC rate of the selected journal. TUD has green

open access to several journals, and in cases when TUD staff are co-authors of the publication, these costs will not apply. The US and UPV/EHU have open access to some journals in the power system field. Therefore, the publication costs do not apply when the US or UPV/EHU researchers are co-authors of the publication. ETF has allocated funds for Gold Open Access in journals.
- 4TU.DataResearch will be responsible for the long-term preservation of the data. TUD is offering access to all project partners.

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

We will not use sensitive data. We will use the default security measures defined by the project board. In this case, only team members can access the designated server. The storage security is ensured by the project coordinator.
Each beneficiary should ensure internal storage and appropriate security measures:

- ETF will use the institutional, personal accounts and store data in the ETF MS Teams.
- TUD will use Surfdrive for internal sharing and for the storage of the consortium files.
- US researchers will use the institutional, personal accounts provided by Microsoft.
- UPV/EHU will use the UPV/EHU MS team's SharePoint folder and the internalGorde Talde tool

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

This project will not lead to the creation of any personal data, and therefore, this question does not apply.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

All consortium partners will store data following 4TU.ResearchData platform upon the completion of the project.
At UPV/EHU, data will be stored internally usingGorde Talde platform.