TradeRES

New Markets Design & Models for 100% Renewable Power Systems

D7.3 - TradeRES Data Management Plan DMP Report

Deliverable number:	*D7.3*
Work Package:	*WP7*
Lead Beneficiary:	*ISEP*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864276

Author(s) information (alphabetical)		
Name	**Organisation**	**Email**
Débora de São José	ISEP	drj@isep.ipp.pt
Filipe Sousa	ISEP	ffeso@isep.ipp.pt
Tiago Pinto	ISEP	tcp@isep.ipp.pt

Acknowledgements/Contributions		
Name	**Organisation**	**Email**

Document information			
Version	**Date**	**Dissemina-tion Level**	**Description**
1.0	28.07.2020	Confidential	Detailed plan describing the data manage-ment procedures

Review and approval		
Prepared by	**Reviewed by**	**Approved by**
Débora São José (ISEP) Filipe Sousa (ISEP) Tiago Pinto (LNEG)	Nicolas Dintzner (TUDelft)	Ana Estanqueiro (LNEG)

Disclaimer

Executive Summary

The present deliverable, as part of task 7.3, addresses the data management plan developed by the consortium members, which describes how the data used and generated by the project will be managed. This report includes the guidelines for how the data are supposed to be handled during the project, aiming its well management, and after its conclusion, preparing for its preservation, sharing and re-use (also focusing on dissemination).

In addition, the data management plan here reported, includes the informed consent procedures as required by the ESR (Early Stade Research).

In what concerns the data to be generated and collected, most of it will be collected at OMIE (Operador del Mercado Ibérico de Electricidad – Iberian Electricity Market Operator), REN (Redes Elétricas Nacionais - National Electrical Networks) and National GRID websites, among other sources. Data generation is expected to be minimal, as described in this document.

Table of Contents

List of Tables

List of Figures

No table of figures entries found.

1. Introduction

This Data Management Plan (DMP) summarizes how the data will be addressed during the project and after its conclusion. This report addresses all aspects of data manage-ment, generation, preservation, storage and analysis. This document also addresses the mecha-nisms that will be used at the end of the project to share and preserve the data.

In that way, this report represents the data management life cycle for the data to be col-lected, processed and/or generated by the TradeRES project.

Based on that, this DMP addresses:

- the handling of research data during and after the end of the project;
- what data will be collected, processed and/or generated;
- which methodology & standards will be applied;
- whether data will be shared/made open access; and
- how data will be curated & preserved (including after the end of the project).

Considering that this is a 2020 Horizon project, it will work with FAIR data, which means that data must be:

- **F**indable, including provisions for metadata: convention names, provide key-words for optimize possibilities for re-use, clear version numbers, among others;
- **A**ccessible: describe data that will be openly accessible and how, the ones that must be shared with restrictions and the ones that cannot be shared;
- **I**nteroperable: allow data exchange and re-use between researchers, institutions, organisations, countries, etc; and
- **R**e-use (increased): licenses to allow data to be re-used and specification about when it will be available for re-use and also for how long it will remain available for it.

In the next sections all pointed aspects will be addressed in the level of detail that is possible, considering that the project is in its initial stage and there is still a lack of detailed information about the data that will be used, as it is dependent on the development of the technical workpackages of the project.

2. Information about data and data format

When planning the data management and considering the data that will/is expected to be used during this project and after its conclusion, the characteristics of the data are crucial to ensure FAIR data, and its accessibility is one of the main aspects for a data management plan.

In what concerns the data to be used during this project, Table 1 summarizes its main characteristics related to Accessibility and Interoperability.

Table 1: General data characteristics

Name	Type / Format	Origin	Expected Size	Utility
MIBEL Day-ahead Market data	Microsoft Excel	OMIE website	1GB	project partners, stakeholders that will experiment/validate the project platform
MIBEL Intraday Market data	Microsoft Excel	OMIE website	1GB	project partners, stakeholders that will experiment/validate the project platform
MIBEL Bilateral contracts data	Microsoft Excel	OMIE website	100 MB	project partners, stakeholders that will experiment/validate the project platform
Energy consumption	SQL database	To Be Defined (TBD)	TBD	project partners
EV behaviour / driving patterns	SQL database	TBD	TBD	project partners
Capacity factors	SQL database	TBD	TBD	project partners
Hydro inflows	SQL database	TBD	TBD	project partners
Power network data	SQL database	TBD	TBD	project partners
Scenario data	SQL database	TBD	TBD	project partners

Portuguese Energy generation and consumption	Microsoft Excel	REN website	400 MB	project partners, stakeholders that will experiment/validate the project platform
Portuguese Secondary Reserve data	Microsoft Excel	REN website	500 MB	project partners, stakeholders that will experiment/validate the project platform
Portuguese Tertiary Reserve data	Microsoft Excel	REN website	500 MB	project partners, stakeholders that will experiment/validate the project platform
Portuguese Imbalances Data	Microsoft Excel	REN website	500 MB	project partners, stakeholders that will experiment/validate the project platform
Iberian Interconnection Exchange data	Microsoft Excel	REN website	100 MB	project partners, stakeholders that will experiment/validate the project platform
Electricity and gas consumption by sector in the UK	Microsoft Excel	National Grid (UK TSO)	1-2 MB	project partners, stakeholders that will experiment/validate the project platform
Electricity and gas peak demand in the UK	Microsoft Excel	National Grid (UK TSO)	100 kb	project partners, stakeholders that will experiment/validate the project platform
Generation capacity in the UK	Microsoft Excel	National Grid (UK TSO)	1-2 MB	project partners, stakeholders that will experiment/validate the project platform

COMPETES model input/output data	Microsoft Access and Excel	Various, including ENTSO-E and IEA databases	500 MB	Project partners and stakeholders
German day-ahead market data	Microsoft Excel	EPEX	1 GB	project partners
Central Europe day-ahead price	Microsoft Excel	smard.de	10 MB	project partners
RE feed-in DE	Microsoft Excel	ENTSO-E	100 MB	project partners
Demand DE	Microsoft Excel	ENTSO-E	10 MB	project partners
Installed power plants DE	Microsoft Excel	open-power-system-data.org	10 MB	project partners
Installed RE capacities DE	Microsoft Excel	www.erneuerbare-energien.de; marktstammdatenregister	10 MB	project partners
Run-of-River generation DE	Microsoft Excel	smard.de	10 MB	project partners
Remuneration data DE	TBD	TBD	1 MB	project partners
Financial Data Renewables	TBD	TBD	1 MB	project partners
CO2 data	TBD	TBD	1 MB	project partners
Fuel cost data	TBD	TBD	10 MB	project partners
Electricity Price	.csv	generated	1 MB	project partners
Revenues	.csv	generated	<1MB	project partners

From Table 1 it is possible to see that some data cannot be completely described at this time, due to some restrictions in the research which do not allow the completely characterization of some specific data. However, during the project, this information will be updated by project partners.

The total storage space used during this project should not exceed 100GB, which helps deciding how to accommodate data. Another aspect is the preference for “easy to access” format/type option.

3. Metadata content and format

The bibliographic metadata must be in a standard format and must include all of the following:

- the terms “European Union (EU)” and “Horizon 2020”;
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable;
- a persistent identifier, and
- dataset authors.

For all archive datasets, the DublinCore and Data Cite metadata format will be followed. This will be achieved by using research data repositories (either 4TU Center for Research Data, Zenodo, IEEE repository, or another suitable repository guaranteeing 10 years of data retention and provides DOI for each dataset).

Each dataset will be documented, either as part of the data itself (additional worksheet in Excel), or in a separate file (README). The documentation will describe the content of the dataset, as well as how the information is organized. Industry standards, formats and units will be used when relevant.

4. Policies for access, sharing and re-use

All sensitive data gathered during the project lifetime will be kept confidential by the consortium, namely electric grid, generation and customers' data. Particular attention shall be paid to personal data which will be codified and destroyed after the project ends. Data protection and confidentiality is also presented in this topic.

4.1 Rights and obligations related to background and results

4.1.1. Management of intellectual property

According to the Article 23a from the Grant Agreement [1], the beneficiaries have the obligation to take measures to implement the Commission Recommendation on the management of intellectual property in knowledge transfer activities, which means that beneficiaries such as universities or other public research organisations must take measures to implement the principles set out in Points 1 and 2 of the Code of Practice annexed to the Commission Recommendation on the management of intellectual property in knowledge transfer activities.

4.1.2. Rights and obligations related to background

According to the Article 24 from the Grant Agreement [1], the beneficiaries must identify and agree (in writing) on the background for the action ('agreement on background'). 'Background' means any data, know-how or information — whatever its form or nature (tangible or intangible), including any rights such as intellectual property rights — that: (a) is held by the beneficiaries before they acceded to the Agreement, and (b) is needed to implement the action or exploit the results.

4.1.3. Access rights to background

According to Article 25 from Grant Agreement, to exercise access rights, this must first be requested in writing ('request for access'). 'Access rights' means rights to use results or background under the terms and conditions laid down in the Grant Agreement.

The beneficiaries must give each other access — on a royalty-free basis — to background needed to implement their own tasks under the action, unless the beneficiary that holds the background has — before acceding to the Agreement —: (a) informed the other beneficiaries that access to its background is subject to legal restrictions or limits, including those imposed by the rights of third parties (including personnel), or (b) agreed with the other beneficiaries that access would not be on a royalty-free basis.

Furthermore, the beneficiaries also must give each other access — under fair and reasonable conditions — to background needed for exploiting their own results, unless the beneficiary that holds the background has — before acceding to the Agreement — informed the other beneficiaries that access to its background is subject to legal restrictions or limits, including those imposed by the rights of third parties (including personnel).

4.2 Rights and obligations related to results

4.2.1. Ownership of results

Results (tangible or intangible output of the action such as data, knowledge or information) are owned by the beneficiary that generates them.

Furthermore, two or more beneficiaries can simultaneously own results if: (a) they have jointly generated them and (b) it is not possible to: (i) establish the respective contribution of each beneficiary, or (ii) separate them for the purpose of applying for, obtaining or maintaining their protection.

4.2.2. Rights of third parties (including personnel)

If third parties (including personnel) may claim rights to the results, the beneficiary concerned must ensure that it complies with its obligations under the Agreement.

If a third party generates results, the beneficiary concerned must obtain all necessary rights (transfer, licences or other) from the third party, in order to be able to respect its obligations as if those results were generated by the beneficiary itself.

If obtaining the rights is impossible, the beneficiary must refrain from using the third party to generate the results.

4.2.3. Licensing

All datasets, which may be publicly released, will be released under a permissive license, such as Creative Commons (CC). We will consider CC-BY to ensure citation.

The software solutions produced during the project that will be shared, should be shared under a MIT (Massachusetts Institute of Technology) or Apache v2 Licenses.

4.3 Protection of results

4.3.1. Obligation to protect the results

Each beneficiary must examine the possibility of protecting its results and must adequately protect them — for an appropriate period and with appropriate territorial coverage — if: (a) the results can reasonably be expected to be commercially or industrially exploited and (b) protecting them is possible, reasonable and justified (given the circumstances).

When deciding on protection, the beneficiary must consider its own legitimate interests and the legitimate interests (especially commercial) of other beneficiaries.

4.3.2. Agency ownership, to protect the results

If a beneficiary intends not to protect its results, to stop protecting them or not seek an extension of protection, the Innovation and Networks Executive Agency (INEA) may — under certain condition — assume ownership to ensure their (continued) protection.

4.4 Exploitation of results

4.4.1. Obligation to exploit the results

Each beneficiary takes measures aiming to ensure 'exploitation' of its results (either directly or indirectly, in particular through transfer or licensing by:

(a) using them in further research activities (outside the action);
(b) developing, creating or marketing a product or process;
(c) creating and providing a service; or
(d) using them in standardisation activities.

4.4.2. Obligation to disseminate results

Unless it goes against their legitimate interests, each beneficiary must — as soon as possible — 'disseminate' its results by disclosing them to the public by appropriate means (other than those resulting from protecting or exploiting the results), including in scientific publications (in any medium).

4.4.3. Open access to scientific publications

Each beneficiary must ensure open access (free of charge online access for any user) to all peer-reviewed scientific publications relating to its results.

(a) as soon as possible and at the latest on publication, deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications;
Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications.
(b) ensure open access to the deposited publication — via the repository — at the latest:
    a. on publication, if an electronic version is available for free via the publisher, or
    b. within six months of publication (twelve months for publications in the social sciences and humanities) in any other case.
(c) ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication.
The bibliographic metadata must be in a standard format and must include all of the following:

- the terms "European Union (EU)" and "Horizon 2020";
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable, and
- a persistent identifier.

4.4.4. Open access to research data

Regarding the digital research data generated in the action ('data'), the beneficiaries must:

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user;
(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results, the confidentiality obligations, the security obligations or the obligations to protect personal data, all of which still apply.

4.5 Data Storage during the project

Each beneficiary is responsible for the selection of data storage solution during the project. However, the storage solution must:

- Ensure sufficient data safety and security (which may include encryption for sharing and transferring personal data);
- Automated backup system, either managed by the partner themselves, or their supporting Information and communications technology (ICT) department;
- Sufficient space for the relevant project data managed by the beneficiary;
- Remote access to the relevant project partners, as required.

Each beneficiary is also in charge of determining whether personal data sharing between partners should be subject to further agreements, under GDPR (processing/co-controller agreement).

A common project repository for non-sensitive data will be put in place, to facilitate data exchange among partners – referred to as TradeRES repository in the following sections.

In what concerns the management and storage of data, Accessibility, Interoperability and Re-use, Table 2 summarizes the guidelines for this process.

Table 2: Storage and Data Management

Name	Open access outside the consortium	Data treatment	Data curation and preservation	How data will be used	Data sharing with project partners
MIBEL Day-ahead Market data	Yes	N/A	Stored in TradeRES data repository	analysis phase, models design, experiments/case studies	Yes

MIBEL Intra-day Market data	Yes	N/A	Stored in TradeRES data repository	analysis phase, models design, experiments/case studies	Yes
MIBEL Bilateral contracts data	Yes	N/A	Stored in TradeRES data repository	analysis phase, models design, experiments/case studies	Yes
Energy consumption	hopefully yes	TBD	TBD	case studies in WP2 and WP5	hopefully yes
EV behaviour / driving patterns	hopefully yes	TBD	TBD	case studies in WP2 and WP5	hopefully yes
Capacity factors	hopefully yes	TBD	TBD	case studies in WP2 and WP5	hopefully yes
Hydro inflows	hopefully yes	TBD	TBD	case studies in WP2 and WP5	hopefully yes
Power network data	hopefully yes	TBD	TBD	case studies in WP2 and WP5	hopefully yes
Scenario data	Yes	N/A	deliverable D2.1	case studies in WP2 and WP5	Yes
Portuguese Energy generation and consumption	Yes	N/A	Stored in TradeRES data repository	Validation phase and experiments/case studies	Yes
Portuguese Secondary Reserve data	Yes	N/A	Stored in TradeRES data repositor	Analysis and validation phases, models design, experiments/case studies	Yes
Portuguese Tertiary Reserve data	Yes	N/A	Stored in TradeRES data repositor	Analysis and validation phases, models de-	Yes

				sign, experiments/case studies	
Portuguese Imbalances Data	Yes	N/A	Stored in TradeRES data repositor	Analysis and validation phases, models design, experiments/case studies	Yes
Iberian Interconnection Exchange data	Yes	N/A	Stored in TradeRES data repository	Analysis and validation phases, models design, experiments/case studies	Yes
Electricity and gas consumption by sector in the UK	Yes	N/A	Stored in TradeRES data repository	analysis phase, models design, experiments/case studies	Yes
Electricity and gas peak demand in the UK	Yes	N/A	Stored in TradeRES data repository	analysis phase, models design, experiments/case studies	Yes
Generation capacity in the UK	Yes	N/A	Stored in TradeRES data repository	analysis phase, models design, experiments/case studies	Yes
COMPETES model input/output data	to be checked with COMPETES model owner/partner (i.e. Dutch government/PBL)	N/A	Stored in TNO project SharePoint	Scenario analysis (WP2), market design (WP4) and case study (WP5)	To be checked with COMPETES model owner/partner (i.e. Dutch government/PBL)
German day-ahead market data	no, proprietary	adjust CET/CEST time-jump, remove last day for leap years	stored at DLR	analysis phase, models design	No

Central Europe day-ahead price	no, already accessible	adjust CET/CEST time-jump, remove last day for leap years	stored at TradeRES data repository	analysis phase, models design	Yes
RE feed-in DE	no, already accessible	adjust CET/CEST time-jump, remove last day for leap years, aggregated to hours	stored at TradeRES data repository	all phases	Yes
Demand DE	no, already accessible	adjust CET/CEST time-jump, remove last day for leap years, aggregated to hours	stored at TradeRES data repository	all phases	Yes
Installed power plants DE	no, already accessible	aggregate to years, derive efficiency curves	stored at TradeRES data repository	all phases	Yes
Installed RE capacities DE	no, already accessible	probably rescaling / adjustments	stored at TradeRES data repository	all phases	Yes
Run-of-River generation DE	no, already accessible	adjust CET/CEST time-jump, remove last day for leap years, aggregated to hours	stored at TradeRES data repository	all phases	Yes
Remuneration data DE	TBD	TBD	TBD	TBD	TBD
Financial Data Renewables	TBD	TBD	TBD	TBD	TBD
CO2 data	TBD	TBD	TBD	TBD	TBD
Fuel cost data	TBD	TBD	TBD	TBD	TBD

Electricity Price	Yes	N/A	stored at TradeRES data repository	all phases	Yes
Revenues	Yes	N/A	stored at TradeRES data repository	all phases	Yes

Table 2 summarizes the main aspects of data management and storage. As presented, some data is still without any description, which is due to the same reasons presented in Table 1. During the project, this information will be filled up by research partners.

Furthermore, the majority of the data will be stored at TradeRES data repository, which is not an issue as it is expected to be around 100 GB.

Table 2 also presents the conditions to allow the data to be access outside of the consortium.

4.6 Resources.

We do have readily available the following support structures:

- ICT departments from each institution providing sufficient and safe storage;
- Ethics committee of TUD in case of reviews;
- Privacy and legal teams for data management support;
- Usage of HPC to be defined.

5. Conclusions

This Data Management Plan summarized how the data is expected to be addressed during the project and after its conclusion.

The DMP addressed how to handle research data during and after the end of the project, it specified and described what types of data will be collected, processed and generated, which methodology and standards are going to be applied, whether data will be shared/made open access and how data will be curated and preserved (including after the end of the project).

Moreover, the DMP also defines how to make the data FAIR: findable, accessible, interoperable and reusable, and the conditions for that to happen during the project and after its conclusion.

Despite some information is still not available in full, as data usage and generation depend deeply on the development of the technical work packages, this report will be updated throughout the project, when precise information on each piece of data becomes available.

References

[1] The Grant Agreement nº 864276 and all its annexes — TradeRES

[2] 864276 TradeRES Consortium Agreement, version 2, 2019-12-12 (drafted from DESCA V1.2.4)

[3] Data Management, available in https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template