



BRESAER

D1.15 2nd version of the Data Management Plan

01/2017

D1.15 2nd version of the Data Management Plan
WP 1, T 1.3

Authors: Karine Laffont (TECHNOFI)

BREakthrough Solutions for Adaptable Envelopes in building Refurbishment
EeB-02-2014 RIA



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 637186.

Technical References

Project Acronym	BRESAER
Project Title	BREakthrough Solutions for Adaptable Envelopes in building Refurbishment
Project Coordinator	Isabel Lacave Azpeitia ACCIONA INFRAESTRUCTURAS mailto:isabel.lacave.azpeitia.ext@acciona.com
Project Duration	February 2015 – July 2019 (54 months)

Deliverable No.	D1.15
Dissemination level ¹	PU
Work Package	WP 1 - Management
Task	T 1.3 Data Management Plan
Lead beneficiary	14- TECHNOFI
Contributing beneficiary(ies)	All
Due date of deliverable	31 January 2017
Actual submission date	31 January 2017

¹ PU = Public

PP = Restricted to other programme participants (including the Commission Services)

RE = Restricted to a group specified by the consortium (including the Commission Services)

CO = Confidential, only for members of the consortium (including the Commission Services)

Document history			
V	Date	Beneficiary	Author
1.0	06/07/15	TECHNOFI	Karine Laffont
2.0	09/12/16	TECHNOFI	Karine Laffont
review	05/01/2017	ACCIONA	Isabel Lacave
Final	20/01/2017	TECHNOFI	Karine Laffont



0 Summary

This deliverable presents the second version of the Data Management Plan (DMP) and has been produced at M24.

The purpose of the DMP is to provide an analysis of the main elements of the data management policy that will be used by the consortium with regard to all the datasets that will be generated by the project. The DMP is a document outlining how research data will be handled during a research project, and after it is completed. The use of a Data Management Plan is required for projects participating in the Open Research Data Pilot and is recommended for almost any other research project.

The DMP is therefore closely related to the Dissemination Plan.

The BRESAER partners will generate various datasets during the project.

Most of these datasets are related to the technologies that are going to be developed. This raises confidentiality issues: indeed, disclosing too much information would open the door to reverse-engineering from competitors. Additionally, if the project results are to be patented, they should not be published beforehand.

On the other hand, it is in the interest of the partners to disseminate the right amount of data about the performances of the technologies (simulation data, data from the demonstration) in order to maximise the exploitation potential.

A compromise must therefore be found between complete confidentiality, partial publication and Open Research Data, and should also be in line with the Dissemination Plan.



Table of content

0	SUMMARY	3
1	OBJECTIVE	5
2	BACKGROUND	5
3	UPDATING THE DMP	5
4	SECOND VERSION OF THE DATA MANAGEMENT PLAN	6
5	DATA SHARING POLICY	12
6	CONCLUSIONS	12



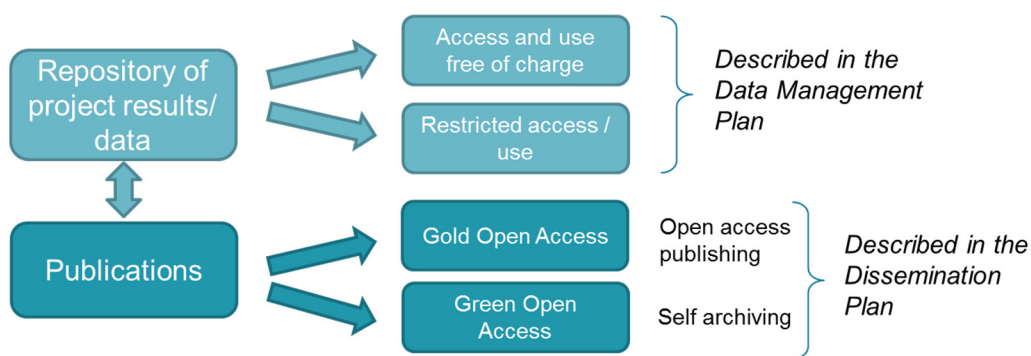
1 Objective

This deliverable presents the second version of the Data Management Plan (DMP) and has been produced at M24.

The purpose of the DMP is to provide an analysis of the main elements of the data management policy that will be used by the consortium with regard to all the datasets that will be generated by the project.

The DMP is a document outlining how research data will be handled during a research project, and after it is completed. It is very important in all aspects for projects participating in the Horizon 2020 Open Research Data Pilot as well as for almost any other research project.

The DMP is closely related to the Dissemination Plan, as pictured below:



2 Background

Data Management Plans (DMPs) have been introduced in the Horizon 2020 Work Programme for 2014-15: *A further new element in Horizon 2020 is the use of Data Management Plans (DMPs) detailing what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. The use of a Data Management Plan is required for projects participating in the Open Research Data Pilot. Other projects are invited to submit a Data Management Plan if relevant for their planned research.*

Projects taking part in the Pilot on Open Research Data are required to provide a first version of the DMP as an early deliverable within the first six months of the project. **Projects participating in the pilot as well as projects who submit a DMP on a voluntary basis because it is relevant to their research should ensure that this deliverable is mentioned in the proposal.** Since DMPs are expected to mature during the project, more developed versions of the plan can be included as additional deliverables at later stages. The purpose of the DMP is to support the data management life cycle for all data that will be collected, processed or generated by the project.

3 Updating the DMP

A DMP describes the data management life cycle for all data sets that will be collected, processed or generated by the research project. It is a document outlining how research data will be handled during a research project, and even after the project is completed, describing what data will be collected, processed or generated and following what methodology and standards, whether and how this data will be shared



and/or made open, and how it will be curated and preserved. The DMP is not a fixed document; it evolves and gains more precision and substance during the lifespan of the project.

According to the EC guidelines, the DMP need to be updated at least by the mid-term and final review to fine-tune it to the data generated and the uses identified by the consortium since not all data or potential uses are clear from the start.

The present deliverable is the mid-term update of the DMP. The final review will be produced on M54 and described in D1.16

4 Second version of the Data Management Plan

The 2nd DMP reflects the current status of reflection within the consortium about the data that will be produced.

The points below will be addressed on a dataset by dataset basis:

- **Data set reference and name**

Identifier for the data set to be produced. (For now only a name is provided. Once the datasets are published/archived, a definitive identifier will be given)

- **Data set description**

Description of the data that will be generated or collected, its origin (in case it is collected), nature and scale and to whom it could be useful, and whether it underpins a scientific publication. Information on the existence (or not) of similar data and the possibilities for integration and reuse.

- **Standards and metadata**

Reference to existing suitable standards of the discipline. If these do not exist, an outline on how and what metadata will be created.

- **Data sharing**

In case the dataset cannot be shared, the reasons for this should be mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).

In the present version of the DMP, since most of the datasets have not been produced yet, two items related to data sharing are described:

- Can the dataset be shared ? (e.g. are there barriers related to confidentiality, privacy, rules of personal data, etc.)
- Can the dataset be re-used within and/or outside the consortium? Only data which underpins published research findings and/or has longer-term value (i.e. can be reused) should be shared.

For the datasets that can be shared and re-used, access procedures will be finalised in the next version of the DMP.

- **Archiving and preservation (including storage and backup)**

Description of the procedures that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered.

The list of datasets and their description will be updated in the course of the project.



Data set reference and name	Task	Partner in charge	Data set description	Date of finalisation of data	Standards, format	Can this dataset be shared?	Is this dataset reusable?	Archiving and preservation (including storage and backup) as foreseen today
All weather 240315.xlsm	T2.1	TNO	Collection of heating and cooling degree days information for 109 locations across Europe. This data was sourced from degreedays.net.	M3	.xlsm	YES	YES	Data stored on project folder in TNO network. Also distributed to other project partners involved in Task 2.1.
Database for Geocluster maps	T2.5	TNO	Database / tabulated data for various parameters (such as climate, building stock typology...) for regions across the EU and Turkey	M36		Can be shared in principle, if no confidential sources are used (in which case there would be restrictions)	YES	At this stage, within TNO servers. May eventually move to another location depending on final host.
Preliminary simulations results	T2.2	Technion	Data containing energy calculations for basecases defined in T2.2 and energy strategy application (results only!)	M11	.csv and xls (EnergyPlus result files)	Will be partially published in a peer-reviewed paper as synoptic charts. If raw results placed on open database server there must be copyright restrictions	Could be reused for virtual demonstrations.	Initial raw data and data analysis to be preserved by partners at least during duration of the project and 5 years later (for audit), but it can have long term preservation by partners performing energy simulations in their off-line backup devices (ie DVDs, USB storage, private server storage, etc). Volume is 1.16 GB.
Simulation results provided by software tool (energy calculations)	T2.3	Technion	Data containing energy calculations generated by tool to be developed in T2.3 using database approach	M54	Metadata to be defined (mysql database)	To be confirmed	Yes with copyright limitations	Data provided by web application that can be downloaded in the final user computers (such as in a computer program) and resides in an online server that provides the program to the final users.



								Costs and maintenance of a suitable web server have to be discussed. Volume not known at this time.
Simulation results provided by software tool (envelope installation)	T4.4	Technion	Data containing envelope installation aids generated by tool to be developed in T4.4	M54	Metadata to be defined (mysql database and possibly BIM files)	To be confirmed	Yes with copyright limitations	Data provided by web application that can be downloaded in the final user computers (such as in a computer program) and resides in an online server that provides the program to the final users. Costs and maintenance of a suitable web server have to be discussed. Volume not known at this time.
Analysis and evaluation of the monitored results	T6.2 T6.7 T6.8	Technion	Energy calculations and other information about the demonstration building (MS4, T6.2, T6.7 and T6.8)	M25 (expected)	Csv and xls (trnsys result files)	Related to WP6 MS4 and D6.3 (Public) Could be used as part of a peer-reviewed paper as synoptic chart If raw results placed on open database server there must be copyright restrictions	In principle yes Part of virtual demonstrations	Initial raw data and data analysis to be preserved by partners at least during duration of the project and 5 years later (for audit), but it can have long term preservation by partners performing energy simulations in their off-line backup devices (ie DVDs, USB storage, private server storage, etc). Volume not known at this stage.
Building information	T6.1 T5.3	CARTIF	Those data related to the measurements of the building and the monitoring network.	M54	LonWorks and IFC4 standards based	Raw data cannot be provided because of upcoming EU regulation (rules of	YES, for virtual demonstration	The data will be persistently stored in database and secure backups will be automatically and weekly generated in order to avoid the loss of data.



					PosgreSQL database	personal data). KPIs only For the public: decision to be taken by the building owner		Additionally, data logs will be maintained to avoid data gaps. This information is easily restored in PostgreSQL database through the backup file.
Technologies data	T5.3	CARTIF	Data collected from the façade solution technologies for the application of the BEMS control algorithms.	M54	LonWorks standard based whenever possible. PosgreSQL database	Open to the consortium. For the public: decision to be taken by the technology owner	YES - by the technology owners only	The data will be persistently stored in database and secure backups will be automatically and weekly generated in order to avoid the loss of data. Additionally, data logs will be maintained to avoid data gaps. This information is easily restored in PostgreSQL database through the backup file.
BEMS data	T5.3	CARTIF	Data generated by the BEMS itself: alarms about malfunctioning, calculation results for the optimization and internal data for rendering the calculations.	M54	data model based on IFC4 for the internal performance of the BEMS and its results. PosgreSQL database	Open to the consortium. KPI-related data will be shared as open data (only about performance)	YES - by the technology owners only, with the exception of aggregated data about performance	The data will be persistently stored in database and secure backups will be automatically and weekly generated in order to avoid the loss of data. This information is easily restored in PostgreSQL database through the backup file.
EMI TEST REPORT	T3.7	Mondragon	Reports and analysis associated with photovoltaic module	M27		See later. If successful, to be used as a marketing support for BRESAER	YES (by Mondragon)	Confidential storage by Mondragon



EMI TEST REPORT	T3.7	Solarwall	Preparation of the material necessary to carry out different tests of the Solarwall material by EMI	M24	ACCORDING TO THE APPLICATION RULES	See later	Yes (by Solarwall)	Confidential storage by Solarwall
EMI TEST REPORT	T3.7	STAM	Reports on tests performed on lightweight insulating panels coupled with and without photovoltaic modules.	M27	ETAG034 Results provided in .doc and .pdf	Relevant data will be disclosed for marketing purposes	YES (by STAM for commercial purposes)	Internal storage by STAM, marketing results will be disclosed through websites
EMI TEST REPORT	T3.7	EURECAT	Reports associated to the test done to the automatic insulated blind. Wind test, thermal test and reaction to fire test.	M27		See later. If successful, to be used as a marketing support for BRESAER	YES (by Eurecat)	Confidential storage by Eurecat
Life cycle analysis and life cycle cost data	T6.4	Tecnalia	Type and quantity of material, cost of material, consumption of energy to manufacturing, description of the production process, ...	M47	ISO 14.040, ISO 14025, ISO 15804.	NO (commercial) See more details below	YES (by consortium) See more details below	Tecnalia will store the data until the end of the project.
Life cycle analysis and life cycle cost data	T6.4	Mondragon	LCC-LCA analysis of polymer concrete ventilated facade module Example: Type and quantity of material, cost of material, consumption of energy to manufacturing, description of the production process	M51		See later. If good, to be used as a marketing support for BRESAER	YES (by Mondragon)	Confidential storage by Mondragon
Life cycle analysis and life cycle cost data	T6.4	STAM	Analysis of costs and environmental impact of production process for the integrated solution of insulating panels + PV elements. Raw materials working procedures and energy consumption are taken into account.	M51	ISO14040 and ISO14044 Results provided in .xlsx (numerical results), reports in .doc and .pdf	Relevant data will be disclosed for marketing purposes	YES (by STAM for commercial purposes)	Internal storage by STAM, marketing results will be disclosed through websites



Life cycle analysis and life cycle cost data	T6.4	Solarwall	Analysis of the life cycle of the solar system that includes not only the Solarwall panel but also the structures. For example: quantity of material, energy consumed, manufacturing, etc.	M28		See Later	Yes (by Solarwall)	Confidential Storage by Solarwall
Life cycle analysis and life cycle cost data	T6.4	EURECAT	LCC-LCA analysis of automatic insulated blind. Example: Type and quantity of material, cost of material, consumption of energy to manufacturing, description of the production process	M51		See later. If good, to be used as a marketing support for BRESAER	YES (by Eurecat)	Confidential storage by Eurecat
Data related to BRESAER substructure	WP3	Mondragon	The result of the project can provide a new kind of profile or substructure that may be patented or protected. The system will also generate data, knowledge and information.	M54		See later. Could be used as a marketing support for BRESAER	YES (by Mondragon)	Confidential storage by Mondragon

(*) Publications:

Technion has budget assigned for one Gold Open Access publication. There might be another publication related to BRESAER, but this would be under the usual copyright agreement of the editorial houses, that would not entail additional charges to the project.

The Gold Open Access publication is likely to be submitted to one of Elsevier's or Taylor & Francis' publications, which offer a large variety of journals with high impact factor that support gold open access publishing. The definitive journal will be selected based partially on budget and the most advantageous open access agreement.

The topic is likely to be based on Task 2.2 (simulations) and Task 2.3 (design tool).



5 Data sharing policy

Most of the datasets generated by the project are related to the technologies that are developed in the project. This raises confidentiality issues: disclosing too much information would indeed open the door to reverse-engineering by competitors. Additionally, if the project results are to be patented, they should not be published beforehand.

On the other hand, it is in the interest of the partners to disseminate a certain amount of data about the performances of the technologies (simulation data, data from the demonstration) to maximise the exploitation potential.

Datasets that have been collected by partners to perform the analyses such as LCA, Geocluster maps, etc., and are not specific to BRESAER technologies, could also be shared with other similar projects.

A compromise must therefore be found between complete confidentiality, partial publication and Open Research Data. The data sharing strategy is at present provisional and will be refined once the datasets are collected/ generated. Once the strategy is finalised, the DMP will describe how data will be (or have been) shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. The repository where data is stored will also be identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).

The consortium is at present investigating the opportunity to use the repository suggested by the EC (<https://www.zenodo.org>).

6 Conclusions

The BRESAER partners will generate various datasets during the project. Most of them are related to BRESAER technologies, which raises confidentiality issues. But datasets which underpin published research findings and/or have longer-term value (i.e. could be reused by other consortia) will be shared, under conditions that will be presented in the final version of the DMP (D1.16, M54).

