D6.2

Data Management Plan (DMP)

Project number:	731453
Project acronym:	VESSEDIA
Project title:	Verification engineering of safety and security critical dynamic industrial applications
Start date of the project:	1st January, 2017
Duration:	36 months
Programme:	H2020-DS-2016-2017

Deliverable type:	ORDP
Deliverable reference number:	DS-01-731453 / D6.2/ 1.0
Work package contributing to the deliverable:	WP6
Due date:	June - M06
Actual submission date:	30 June 2017

Responsible organisation:	CEA
Editor:	Armand Puccetti
Dissemination level:	PU
Revision:	1.0

Abstract:	This document constitutes the Data Management Plan (DMP) of the VESSEDIA project, explaining how the project plans to manage publications and research data.
Keywords:	Data Management

The project VESSEDIA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731453.

Editor

Armand Puccetti (CEA)

Contributors (ordered according to beneficiary numbers)

Shuai Li, Boutheina Bannour, Mounir Kelil (CEA)

Duquesnoy Simon (INRIA)

Marion Habernig (TEC)

Dillon Pariente (DA)

Vendel Laszlo (SLAB)

Cedric Berthion (AMO)

Jacobs Bart (KUL)

Jens Gerlach (FOKUS)

Igor Grueiro Santos (DE)

Paakkola Roosa (TUAS)

Executive Summary

This document constitutes the Data Management Plan (DMP) of the VESSEDIA project, describing the data management life cycle for the data to be collected, processed and/or generated by this Horizon 2020 project. The overall aims are to make research data findable, accessible, interoperable and re-usable (FAIR).

This document has been compiled as a summary of a questionnaire-based survey, according to the EC Guidelines on FAIR Data Management in Horizon 2020, which was distributed to all the partners of the VESSEDIA consortium.

The partners' questionnaires show that the project generates quite some data whose characteristics are partially known. Indeed, since it is quite early in the project, volume, structure, sharing, archiving, and licencing are partially known, even though the project is doing its best to open-source it. The project is aware of these aspects and will update the present document when they become precise during the development of the specifications of the experimentations. The management architecture (SVN, website, etc.) is well defined central Data management structure and will prove its efficiency along the project, to manage data and share it between the partners as well as with other entities involved in the project (reviewers, project officer, advisory board, etc.).

Notice that Data are subject to all kinds of changes, given that this project is about research. Therefore, information in this document is subject to change and updates will be included in the periodic reports.

Contents

Chapter 1 Introduction

The H2020 programme has implemented a pilot action on open access to research data. VESSEDIA, as a participating project to this pilot action, is required to develop a **Data Management Plan (DMP)**. This DMP has been identified in the Description of Action (DoA) as VESSEDIA deliverable D6.2. This document is drafted according to the **"Horizon 2020 FAIR Data Management Plan template"**. The major aim of the DMP is to ensure that our research data is FAIR – **F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable.

Thus, by the DMP we aim to address data set references and names, data set descriptions, standards and metadata, data sharing and archiving and preservation (including storage and backup) on a dataset by dataset basis.

The DMP is intended to be a **living document**. It will be periodically revised to reflect changes in the data that may be made available by the project, and to provide additional information on the datasets as this information is developed during the specifications of the experimental phases.

All partners have contributed to the document, particularly through the use of a project wide questionnaire.

Since each partner will generate and use data, the document is organized with one section per partner. Each section is structured following the **6-points structure** described thereafter:

1. **Data summary** gives a description of the data.
2. **FAIR data** ensures that data are **F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable.
    - 2.1 **Making data findable, including provisions for metadata** gives a description of the data.
    - 2.2 **Making data openly accessible** explains how the data are made openly accessible.
    - 2.3 **Making data interoperable** assesses the interoperability of the data.
    - 2.4 **Increase data re-use (through clarifying licences)** discusses how the re-usability of the data is ensured.
3. **Allocation of resources** refers to costs and other resources for making the data FAIR.
4. **Data security** addresses data recovery as well as secure storage and transfer of sensitive data.
5. **Ethical aspects** clarifies if there any ethical or legal issues that can have an impact on data sharing.

In the following, we will address the methodology that was used to set up the deliverable. Furthermore, we will provide a Data summary, followed by a summary of the use of FAIR data in the VESSEDIA project. Finally, allocation of resources, management of data security, as well as ethical aspects will be elaborated. .

Chapter 2 Methodology

In order to compile the data management plan, a questionnaire was first elaborated, covering the main questions that need to be answered in the template provided by the European Commission.

In a second phase, each project partner responded to the questionnaire, filling it with as much detail as possible at this stage of the project. Completed questionnaires were stored for analysis and traceability in the project's SVN repository.

In a third phase, the Data Management Plan was created as a synthesis of the questionnaire results, attempting to take advantage of commonalities between responses, in order to provide a simple view of data management procedures within the consortium.

Further revisions of the document will be based on updates to partner questionnaires. Therefore, the DMP will be updated at least by the mid-term and final review to be able to fine-tune it to the data generated and the uses identified by the consortium.

In addition, a confidential index of datasets will be created and maintained in the project when the datasets are created. The DMP itself is a confidential document. Therefore the information about the datasets will only be provided to the EU and the reviewers.

The VESSEDIA project will consider open licenses and open availability for the datasets. The reasons for not offering open access will be documented in the partner questionnaires and in the appendix describing the datasets.

Chapter 3 Data Summary

3.1 Generalities

VESSEDIA is a research project that produces mainly data through **Research and software Developments (R&D)**.

The partners will use background R&D data, such as original research articles and other projects' reports and tools, to do research, and produce new data of various kinds, including reports and articles, source and binary code (mainly for x86_64 platforms running Linux, MS Windows and MacOS X, as well as IoT target platforms), archives (containing various kinds of files), log files (containing execution traces and results, proof traces, intermediate results, etc.), text files (containing documentation, traces, etc.), script files (e.g. with compilation, installation and execution procedures), etc.

Various tools will be used and developed by the project:

- Software development tools (commonly called CASE tools),
- Software V&V tools (for the testing and analysis of code and models),
- Modelling tools (for the development of software models),
- Documentation writing tools (such as text editors),
- Configuration management tools (mainly GIT, SVN and Zenodo),
- Workflow management tools (often internal to partner organisations),
- etc.

The second category of tools is the main purpose of the project on which the partners will focus.

The data will be formatted as much as possible using standard format types (and respective naming conventions):

- Reports are mostly in HTML, Markdown, Office Open XML, OpenDocument, and LaTeX format, and therefore files have the corresponding extensions (.docx, .pptx, .xlsx, tex, .html, .md, .pdf, etc.).
- Source files of tools are plain text files using standard naming convention corresponding to the programming languages and computing platforms used: .c, .cpp, .ml, .java, bash, sh, perl, .py, etc. Some text documentation (e.g. README) and script files (e.g. Makefile) are often distributed with these source files (they have no specific extension).

These lists are not exhaustive.

Documentation is produced alongside with research activities, either when new results are found or when planned deliverables need to be produced. Metadata, as stated above, are automatically associated to the tools that serve to prepare the documentation (text editors, MS Office, LibreOffice, LaTeX, etc.), or by the **software development tools** (gcc, MS Visual C++, Ocaml, etc.). The **metadata types** for documentation have been chosen by the project among the most widespread standards. The metadata types for software files are imposed by the development tools used on Linux, MS Windows and MacOS computing platforms.

The size of these kinds of data is not known nor bounded.

The data will be **useful to several parties**:

- The Computer Science community, especially to those interested in Formal Methods and Testing, Embedded systems, and IoT. This includes all possible kinds of organisations, ranging from academia to SME and industries, mainly those producing safety and security critical software.

- The project members as well as other similar research projects members (either national or European) for cross-fertilisation purposes.

Their purpose is manifold and includes:

- Research
- Training: users, students, etc.
- Consultancy
- Applications developments: tools will be used on various kinds of use-cases to demonstrate their effectiveness and usage on real cases
- Tools developments: new and improved tools
- Certification: data is useful for CC evaluation and certification purposes.

3.2 Partners specificities

Below are the specificities of the project partners in terms of data types, collection and generation. Accessibility of data will further be dealt with in section 4.2.

TEC: As the project Coordinator, Technikon will generate data for the Quality Management of the project. This data will contain text, tables, graphs, electronic measurements (e.g. SRAM start-up behaviour) and maybe also some code. Data formats are chosen accordingly, to make the data shareable and accessible in the long term. Data will be reused from scientific publications, conclusions from conference papers, technical data sheets from manufacturers and from other research projects, such as UNIQUE, HECTOR, certMILS, etc.

CEA: The Frama-C V&V platform, the Papyrus Eclipse project, the Diversity tool and the 6LowPan use-case are part of the background material.

DA: DA is building its use-cases and possibly complementary analysis tools. DA will use an internal development methodology for this purpose. Its use-case will be partly available publicly, possibly on the project's website. Confidential parts will not be available.

SLAB: SLAB data consists of reports and evaluations and will be publicly accessible.

FOKUS: FOKUS data consists mainly of reports, specifications, some source code, and scientific articles that are produced by original research activities of the project.

INRIA: The Contiki source code is part background material it is available publicly on Github.

TUAS: The Economic rationale, as a tool for supporting decision making, can be used for a wide range of products/systems' owners/designers/builders, not only applicable to products/systems close in nature to the Vessedia pool of use-cases. So the Economic rationale has a lot of potential for being integrated and re-used.

KU Leuven: The VeriFast V&V tool is part of the background material.

FD: FD will be using the benchmark SPEC for testing efficiency of our tools.

AMO: AMO will use an internal development methodology.

Chapter 4 FAIR Data

In this chapter we describe what is done to make the data FAIR (**F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable).

4.1 Making data findable, including provisions for metadata

4.1.1 Generalities

The project's data is discoverable. As described above, the standards chosen for identifying data correspond to naming conventions (file extensions) at the data files level. Such files are generally compact, easy to manage and can be managed by standard applications, as follows:

- HTML files (suffixes .html, .htm) can be opened by any web browser
- Markdown files (suffix .md) can be handled by any text editor (e.g. emacs)
- Microsoft Office documents (suffixes .docx, .xlsx, .pptx, etc) must be handled by MS Office
- Microsoft Project documents (suffix .mpp) can be edited using MS Project or any compatible tool on Linux
- OpenOffice documents (suffixes .odt, .xml, .rtf, etc.) can be handled by OpenOffice and LibreOffice
- LaTeX format (suffix .tex) must be handled by LaTex
- Acrobat PDF format (suffix .pdf) must be opened by Adobe Acrobat.
- Source files of tools (suffixes .h, .c, .cpp, .ml, .mli, bash, sh, .bat, .perl, .py, etc.) must be handled by the development tools (e.g. MS Visual C++, gnu tools, etc.) or by any text editor (e.g. emacs)
- Assembly files (suffixes .s, .asm) for Intel x86 and ARM architectures mainly that will be handled by assemblers or assembly code analysers.
- Proof script files must be handled by proof tools (such as Qed, Why3, PVS, AltErgo, Z3 and Coq)
- Binary files are mostly the result of the development tool chains (suffixes .exe, .com, .o, .so, .a, .dll, .lib, etc.) and can be handled using some debugger (e.g. gdb) or profiling tools or binary code analysers.

These files naming conventions have the benefit of allowing Operating Systems to identify them easily as well as associating some tool(s) for opening them automatically. This association can be changed easily in each OS.

4.1.2 Partners specificities

Below are the specificities for each project partner:

TEC: None.

CEA: Versions of Frama-C will be named by an element of the periodic table of chemical elements and a version number.

Papyrus follows the Eclipse release train and has a persistent identifier according to the release. Plugins for Papyrus, developed within VESSEDIA, may have their own persistent identifier but it is loosely correlated to the Papyrus identifier.

Diversity's components will be made available at different moments of the projects according to their maturity. They will have an incremental persistent identifier.

DA: Data will be available at the delivery date in the public deliverable reports. No specific identifiers are used.

SLAB: None.

FOKUS: The report on formal methods will have a version number that also refers to the version number of Frama-C.

INRIA: None.

TUAS: Online E- training materials will also be generated by TUAS. Notice that the volume of data of TUAS is hard to assess at this point. For example, the ISO standard documents are large documents (approximately 80 pages).

KU Leuven: Platform/software versions information are used too.

FD: None.

AMO: None.

4.2 Making data openly accessible

4.2.1 Generalities

The major part of the project data is accessible. The project produces whenever possible:

1) Software components with an open-source licence (e.g. LGPL, Berkeley, GPL and MIT), and
2) Public documents (such as deliverables or white papers or thesis reports).

When such components are deemed to be of sufficiently high quality, open source software components can be downloaded on a public web site, on the project web site or on Gitlab or through some Linux packages manager (such as Synaptic on Ubuntu).
Other software components can be made available to partners in prototype state, depending on the author's decision. They may then be made available on the internal and/or external project web site or may be transmitted directly between partners.
Finally, when data is classified as confidential, it will not be available outside of the project.
Data embedded in some public document is public data and becomes visible.
In order to access public data, web browsers are the most suitable. Users may also use, when necessary, some FTP, GIT or SVN client software (e.g. filezilla, rapidsvn, etc.) to transfer the data to a local computer.

In terms of locations, potential users can find the project's public data on the following places:

- Documents, dissemination of reports, open source prototype tools, test cases, white papers and demonstration videos can be found on the project's web site: https://vessedia.eu. Additionally, a contact page of the project website will allow interested persons to contact the project coordination team through some contact form
- In the case of conference papers, they can be obtained as part of conference proceedings, if their organisers decide to let the material be publicly available. Otherwise, they can be obtained within the paper versions of these proceedings.

4.2.2 Partners' specificities

Here are some specificities for each project partner:

TEC: It is possible that some public data will also become available through a thesis developed within the scope of the VESSEDIA project. The thesis may be closed for a few years after its finalization. Any relevant information (research data like SRAM readouts or statistical evaluations) will be shared with the partners internally via the project subversion repository (https://vessedia.technikon.com). Other information, which is also relevant for the global publis will be shared via the project website, e.g. the "results section:" https://vessedia.eu/results.

CEA: Frama-C can be found on the public web site http://www.frama-c.com, in the repositories of the major Linux distributions such as Ubuntu, as well as in the repository of the OCaml package manager (https://opam.ocaml.org).

Papyrus is and Eclipse project and it is available at http://www.eclipse.org/papyrus. The source code of Papyrus is available in an Eclipse repository, available on the Papyrus website.

Diversity is distributed open source as an Eclipse project and can be found at the public website of EFM https://projects.eclipse.org/projects/modeling.efm and associated Eclipse repositories. Sharing other components that are not part of the distribution will be done directly by CEA according to some agreements.

The 6LowPan use-case provided by CEA is confidential and its source code will not be made public. However, it can be transferred to project partners after a NDA has been signed with CEA.
Some further data may come from PhD students, master degree or postdoc theses.

DA: None.

SLAB: None.

FOKUS: Data will be available and useful for the wider public but is not usable for non-specialists.

INRIA: The verified Contiki OS source code will be provided on a dedicated, publicly hosted repository.

TUAS: Data is discoverable and accessible to a certain degree. TUAS will share data with project partners and also with students of their university.

KU Leuven: Versions of VeriFast, together with its documentation will be public and available on the author's server at https://people.cs.kuleuven.be/~bart.jacobs/verifast/

FD: Data produced by FD will be available to people inside Program Analysis field, and more generally to Computer Science Scientists.
AMO: Most AMO data is public but some reports are confidential and will thus not be available outside of the project.

All: To guarantee open access to scientific publications and research data, several repositories were selected (see above). They are convenient to access and also easy to use. Those repositories allow to easily share the long tail of small research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science.

Further, each uploaded publication and dataset receives a persistent identifier (DOI), which ensures the long term preservation.

4.3 Making data interoperable

The project's reports are trivially interoperable as they can be edited and reused with any compatible editor (see above).

In terms of software produced by the project, these are not easily interoperable with other software tools as they generally require the same tool equipment. Software source code can be reused by installing the same development tools as for their production, to read, edit and recompile them. Verification data (predicates, lemma, axioms, proofs, proof scripts, etc.) is generally added to the source code and can only be processed by the same specialised tools as the project (Frama-C, Verifast, Flinder, Why3, VeriFast, etc.) and thus is not interoperable with other tools.

- Frama-C state and script files can only be read by Frama-C.
- Software models are stored in XMI (XML Metadata Interchange) format are also specialised data that can be exchanged with other tools accepting XML. XMI data is compliant to EMF-UML2 [1] that is am implementation of UML2 by EMF.
- Diversity has a proper language and the UML format can be used to exchange data.

Other data in text format such as log or trace files can be edited and reused.

4.4 Increase data re-use (through clarifying licences)

Data reuse is an issue specific to each partner who produces data. Below is described how each partner plans to manage data reuse within the VESSEDIA project:

TEC: Management and project quality data do not need any licence and there is no restriction on the reuse of third-party data, nor any restrictions to share it.
Quality of data produced is ensured by internal checks such as peer reviews.

CEA:

- **Licencing issues**
  - In terms of licensing, an open source (LGPL 2.1) version of Frama-C has been released since the beginning of the VESSEDIA project, and new versions will be published during the course of the project, which will incorporate extensions made within the project. The latest release, version 15 - Phosphorus, has been published on May 31th, 2017. New components that will be developed for Frama-C platform and that will be distributed get an open-source licence whenever possible.
  - Papyrus and all of its plugins are open source and distributed under the Eclipse Public License 1.0 (EPL 1.0). Papyrus implements open-source standards provided by the Object Management Group (OMG). Within VESSEDIA, new plugins developed for Papyrus will continue to be open-source and licensed under EPL 1.0.
  - Diversity is under an open source process as an Eclipse project "Eclipse Formal Modelling Project" (EFM for short), licensed under Eclipse Public License 1.0 (EPL 1.0): A first open source version of Diversity will be realised at the

[1] https://www.eclipse.org/modeling/mdt/?project=uml2

beginning of the VESSEDIA project. New versions with upgrades will be released during the project in the same context. New formal analysis modules and gateways of Diversity that will be developed will be licenced under EPL and distributed through the EFM project whenever possible.
  - The CEA 6LowPan application is confidential and will only be shared within the project between partners that signed a NDA with CEA.
- **Sharing**
  - Frama-C will be shared only when CEA estimates that it is mature enough and by means of new public versions (as for version 15 - Phosphorus). Some new components may be restricted when they are not part of the Frama-C distribution or when they are subject to some agreement with a third party.
  - New plugins developed for Papyrus, within VESSEDIA, will be shared only when CEA estimates that they are mature enough. Some new components may be restricted if they are subject to some agreement with a third party.
  - Diversity will be shared only when CEA estimates that it is mature enough and by means of new public versions (under EFM project, EPL licence). Some new components may be restricted if they are not part of the Diversity distribution or if they are subject to some agreement with a third party.
  - When data is part of some agreement with another company, access will be defined by both parties specifically.
  - Otherwise and if data is not part of the Frama-C distribution, it can be accessed through particular agreements with CEA.
  - In the case of Papyrus, if the data is not part of the Papyrus distribution it can be accessed through particular agreements with CEA.
  - For Diversity: when data is part of some agreement with another company, access will be defined by both parties specifically. Otherwise, if data is not part of the Diversity distribution it can be accessed through particular agreements with CEA.
- **Data quality** is ensured at CEA in several ways: articles and deliverables are project internally and peer reviewed in order to ensure a high quality standard; software components are intensively tested and peer reviewed within the development team (who are assumed to understand the internals of the tools as well as mastering the programming languages used). Furthermore, bug tracking systems and discussion lists ensure that users can report bugs, discuss issues and contact the development team to solve problems. See https://frama-c.com/support.html for instance.

DA: Parts of the use case shared on the public project site, and public reports (project deliverables), unless duly mentioned, have no particular license. Then, there is no restriction on the reuse and sharing of these unlicensed data.

SLAB: Same as for DA.
FOKUS: FOKUS will continue updating a report on formal methods that is licenced under https://creativecommons.org/licenses/by-nc-sa/3.0. There are a few restrictions, see https://creativecommons.org/licenses/by-nc-sa/3.0/

Potential users can find reports and data on the VESSEDIA website and on the website of Fraunhofer FOKUS https://www.fokus.fraunhofer.de. The manual “ACSL by Example” is accessible at https://gitlab.fokus.fraunhofer.de/verification/open-acslbyexample.git

INRIA: Concerning Contiki OK, the new source code will be part of Contiki and will adopt the same BSD license. There is no restriction on the reuse of third-party data, nor any restrictions to share it. New versions are available on the http://www.contiki-os.com main site.

The quality of Contiki is ensured by peer review and checks. New distributions of the OS are intensively tested before release. Bugs are centralized through the Contiki support (at http://www.contiki-os.org/community.html) in order to be processed.

TUAS: ISO/IEC will licence only the final version of the standard document. All the draft versions of ISO/IEC standards will stay available for any expert willing to give comments through a national body or a liaison organisation.

KU Leuven: KUL is using permissive open source license (MIT)/permissive creative commons license.

FD: FD data will be open-source. The precise license has still to be determined.

AMO: AMO reports are public but some reports are confidential and will thus not be available outside of the project. There are no restrictions for the reuse, sharing or access of data.

Chapter 5 Allocation of resources

5.1 Generalities

The costs for making the data FAIR are different for each partner and will be discussed below.

In terms of **data management**, each partner is responsible for handling its internal data:

- The management of data shared in the project, such as on the project's SVN, is done by TEC and the partners. The administrator is TEC, who established rules for partners to use this facility to store, share, modify and query data. Data volumes are not limited so far and will be reasonable.
- Data which is made public on the project's web site has a shared responsibility even though the management of this server is done by TEC.
- Other published data, such as articles, white papers and tools, engage the responsibility of authoring partners as well as the entire project, who approves such data before publication.

5.2 Partners specificities

In terms of **long term preservation**, each partner has a different policy.

TEC: Costs incur for server provision and maintenance (expected costs of about 2000€).
TEC has foreseen to retain the project's data for 3 years after project. The expected costs are those that arise through server provision and maintenance.

CEA: The Frama-C website is hosted on a rented server. Costs for renting the server are part of the lab's operating costs, and are shared among all projects and industrial contracts that are using Frama-C. The same applies to Papyrus. There are no costs for Diversity.

Components of Frama-C and Papyrus that are not released will be kept until mature enough, unless a cooperation agreement is signed with another organisation that makes use of it. Frama-C as well as public plug-ins are archived on the public web site http://www.frama-c.com. Papyrus and its plugins are available at http://www.eclipse.org/papyrus and the source code repository is available on the website.

For Diversity, data that is not released will be kept until mature enough, unless a cooperation agreement is signed with another organisation that makes use of it. Diversity source code is available at the following public repositories http://git.eclipse.org/c/efm/org.eclipse.efm-modeling.git (for the GUI and modelling part) and http://git.eclipse.org/c/efm/org.eclipse.efm-symbex.git (for the formal analysis modules).

In terms of data preservation, tools will be kept for several years (including released versions and tool prototypes). Released tools survive years and prototypes decay unless researchers take up in the scope of a new project. This is the case for Frama-C, Papyrus and Diversity.

DA: There are no extra costs as data is stored on the project's sites.

SLAB: There are no extra costs as data is stored on the project's sites.

FOKUS: No cost as it is hosted by Fraunhofer FOKUS. Maintenance of server is guaranteed by basic funding of Fraunhofer FOKUS. There are no plans to limit the availability.

INRIA: No costs are incurred to make data FAIR.
The data will be publicly hosted on Github on a free account and retained as long as Github provides free public repositories.
Applications specific data as well as non open-source modules will be retained.

TUAS: No costs are incurred to make data FAIR.
TUAS data will be discoverable and accessible to a certain degree. When this is the case, it will be usable beyond the original purpose, e.g. the project's ISO standard will enable building new ISO standards, new tools, etc.

KU Leuven: KUL data will be kept indefinitely. KUL will use Zenodo for storage. We already prepare the data for review by reviewers of the scientific publications. Extra effort for preservation is expected to be minimal.

FD: There are no extra costs as data is stored on the project's sites.

AMO: There are no extra costs as data is stored on the project's sites.

Chapter 6 Data Security

Data security is concerned by the secure handling, storage, loss prevention, recovery and transfer of project data, especially for sensitive data. Data security is managed differently by each partner, as follows:

TEC: Data is secured by saving it regularly on external server (daily), which also avoids losses. Data can be recovered by seeking for the latest copy on external server (max. 24 hours are lost). Sensitive data is handled inside the SVN repositories.

CEA: Frama-C related data is stored on a server which runs a regularly updated Linux distribution, and only accessible through ssh by a few administrators.
Data on the server is part of a git repository that is mirrored in several places, physically distinct from the main server.
Data recovery can be done by simply cloning from one of the mirror repositories.
The 6LowPan application is sensitive data and is stored by CEA internal servers and its security is compliant with the CEA security policies.

DA & AMO: The security of DA publicly available data relies on the project's server managed by TEC. Data is duplicated into DA internal repository, including official deliverables (following ISO 9001 quality management systems - requirements).
Sensitive data will be handled by any means in accordance to the project public web site management.

SLAB: The security of data of SLAB data is built into the project's server managed by TEC.
Data is backed up into an SLAB internal repository.

Sensitive emails are encrypted by PGP. Secure data can be stored in an offline network.

FOKUS: Security will be maintained by Fraunhofer's IT infrastructure. Backups are maintained. GIT- repositories can simply be fully-cloned, making recovery from a clone very easy.

INRIA: The data will be as secure as Github allows. Loss of data is prevented by Github by offering cutting-edge redundancy services. All contributors have copies of the repository on their work station. Data recovery can be done by simply cloning from one of the mirror repositories.
There is no sensitive data.

TUAS: Internal TUAS security measures enable data to be stored securely and avoid any loss of data. Sensitive data will be handled according to effective national and EU-wide privacy regulations.

KU Leuven: Security and recovery procedures are dealt with by Zenodo. There is no sensitive data.

FD: The security of data of FD data is built into the project's server managed by TEC as well as in the Git servers.

Chapter 7 Ethical Aspects

To the best of our knowledge, neither the project nor its later use, nor other planned exploitation of the results have an impact on ethics, but complies with the general social rules of protecting the citizens' right on its own data and purpose of communications. The partners will conform to their current national legislation and regulations.

Chapter 8 Summary and Conclusion

The Data Management Plan of VESSEDIA describes the activities of the partners related to datasets and is a key element of good data management. In the above sections we described the methods used for these activities and provided an extensive definition of the data and their formats. We described how data is made FAIR in the VESSEDIA project as well as the choices made by the partners for the allocation of resources and data security. We have structured all these items into Generalities and Partners Specificities. This document contains a summary of all the information available as of 9th June, 2017. All partners intend to create data and make it available within the consortium.

The DMP needs to be updated in the course of the project whenever significant changes arise, such as new data, changes in consortium policies (e.g. new innovation potential, decision to file for a patent), changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving). This will be done within the periodic reports in M18 and M36.

The partners' questionnaires have helped to understand how each partner uses and manages data. Through the questionnaire, we have also been able to establish that a set of common practices and tools already exist to manage data, allowing date to be shared and reused easily.

Chapter 9 List of Abbreviations

Abbreviation	Translation
BSD	Berkeley Software Distribution
CASE	Computer Aided Software Engineering
DMP	Data Management Plan
DoA	Description of Action
EFM	Eclipse Formal Modelling project
EMF	Eclipse Modelling Framework
EPL	Eclipse Public License
FAIR	Findable, Accessible, Interoperable, and Re-usable
FTP	File Transfer Protocol
Frama-C	Framework for the analysis of C programs
GIT	Version control system
Github	Software hosting system using Git for versions control
Gitlab	Repository management software tool using Git
GPL	GNU Public Licence
GUI	Graphical User Interface
HTML	Hypertext Markup Language
LaTex	Lamport TeX, a language for text processing
LGPL	GNU Lesser General Public License
MS	Microsoft
Ocaml	Objective Categorical Abstract Machine Language
OS	Operating System
PGP	Pretty Good Privacy (encryption software)
SVN	Sub-version control system
UML	Unified Modelling Language
V&V	Verification and Validation
XMI	XML Metadata Interchange
Zenodo	Repository of research publications