

D1.2 Data Management Plan



Funded by
the European Union

D1.2 Data Management Plan

Dissemination Level: Public
 Lead Partner: TU Delft
 Due date: 31 March 2024
 Actual submission date: 31 March 2024

PUBLISHED IN THE FRAMEWORK OF
 MULTICARE (Horizon Europe grant 101123467)

AUTHORS

Alessandra Luna Navarro, TU Delft
 Justin Schembri, TU Delft
 Simona Bianchi, TU Delft

REVISION AND HISTORY CHART

VERSION	DATE	EDITORS	COMMENT
V0.1	04 March 2024	Simona Bianchi (TU Delft)	First Draft
V0.2	15 March 2024	Simona Bianchi (TU Delft) Alessandra Luna Navarro (TU Delft)	Second Draft
V0.3	21 March 2024	Justin Schembri (TU Delft)	Third Draft
V1.0	25 March 2024	Justin Schembri (TU Delft)	Final Version

DISCLAIMER

The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

All rights reserved

The document is proprietary of the MULTICARE consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights. This document reflects only the authors' view. The European Community is not liable for any use that may be made of the information contained herein. Responsibility for the information and views expressed in the therein lies entirely with the author(s).

Executive Summary

The Data Management Plan (DMP) for the MULTICARE project outlines the handling of data throughout its lifecycle, emphasizing retrieval, storage and management during and after the project. It aims to ensure that data is findable, accessible, interoperable, and reusable (FAIR) by providing clear guidance on availability. By adhering to FAIR principles, the data can support the demonstration, adoption and further development of MULTICARE solutions, as well as facilitate ongoing research efforts.

The document offers an overview of the diverse datasets and project outputs, including software, and outlines measures to ensure FAIR management. The DMP is a dynamic document, continuously updated to support ongoing project activities related to data and output management.

Table of contents

D1.2 Data Management Plan	2
PUBLISHED IN THE FRAMEWORK OF	2
AUTHORS	2
REVISION AND HISTORY CHART	2
DISCLAIMER	2
Executive Summary	3
Table of contents	4
LIST OF TABLES	5
GLOSSARY	6
1. Introduction	7
2.1. MULTICARE project	7
2.2. Data Management Plan	9
1.2.1. FAIR principles	9
2. Data Summary	11
2.1 Interviews and workshop data	11
2.2 Existing data	12
2.3 Generated, synthetic and sample data	13
2.4 Collected data	16
2.5 Digital tools	16
2.6 Documentation, Guides and Educational Material	17
2.7 Communication and Dissemination	18
3. FAIR data	19
3.1 Making data Findable, including provisions for metadata	19
3.2 Making data accessible	20
3.3 Making data interoperable	20
3.4 Increase data re-use	21
4. Other Data Management Provisions	22
4.1 Allocation of resources	22
4.2 Data security	22
4.3 Ethics	22
5. Conclusion	23

LIST OF TABLES

Table 1. Consortium	8
Table 2. Interviews and workshop data	11
Table 3. Existing Data	12
Table 4. Generated, synthetic and sample data	13
Table 5. Existing Data	16
Table 6. Digital Tools	17
Table 7. Documentation, Guides and Educational Material	17
Table 8. Communication and Dissemination	18

GLOSSARY

ACRONYM	FULL NAME
API	Application Programming Interface
IEQ	Indoor Environmental Quality
ISO	International Standardisation Organisation
OEQ	Outdoor Environmental Quality
OGC	Open Geospatial Consortium
OMG	Object Management Group
CDE	Common Data Environment

1. Introduction

2.1. MULTICARE project

The built environment is ill-prepared for more frequent and increasingly intense climate-related extreme events. The current building stock is particularly vulnerable because it has limited or no capacity to adapt and recover from extreme events thereby leading to building failures that cause severe socio-economic losses and adversely affecting the health and wellbeing of people. Recent scientific and technological advances in the construction industry provide timely solutions for improving the resilience for specific single hazards (e.g. flood hazard or seismic hazard), but they are often not cost effective, rarely eco-friendly and nearly never address the multiple hazards present in many locations. This is hardly surprising because there is neither a clearly defined framework for quantifying the whole-life socio-economic-environmental impacts of extreme natural events nor tools for assessing the holistic climate resilience of buildings. Consequently, it is currently very challenging to develop/select optimal solutions for real-world multi-hazard scenarios.

MULTICARE will address this challenge directly by developing new multi-criteria decision-support frameworks and providing plug & play technological and digital solutions for improving the resilience of the built environment in a cost-effective, reliable and sustainable manner. The technological solutions consist of multi-functional low-carbon resilient technologies embedded in modular and prefabricated construction for the next generation of high performance and smart buildings, characterized by enhanced safety, energy efficiency, environmental-sustainability, improved quality of life, circularity, and scalability for a broad range of natural events and end-user. The plug & play technologies will be applied to either new multi-story buildings or existing structures by means of low-invasive external interventions. The digital solutions consist of a suite of multi-disciplinary digital services and tools for performing multi-hazard resilience assessment, design, operation and management across multiple scales (material, component, building, neighbourhood/city). The new digital tools will enable stakeholders to make informed decisions in the selection of materials/solutions, including for heritage buildings, and support resilient supply chains. The effectiveness of the MULTICARE solutions will be demonstrated through large-scale pilots (3 buildings, 4 neighbourhoods/district) in three different European countries carefully selected for their diverse local environmental, social and economic conditions (Italy, Netherlands, Romania). Banks and institutional investors will be engaged to better understand the financial risk reduction value of resilience and update existing and future “green finance” mechanisms that will help to leverage the project results. A user-centred, inclusive and participatory approach will be consistently implemented throughout the project to engage citizens and extend the durability of MULTICARE impact.

To achieve these ambitious goals, MULTICARE brings together a unique interdisciplinary Consortium of 21 partners (Error! Reference source not found.) from 6 different EU countries with strong R&D and practical expertise, who are either established leaders in

their sector or agile SMEs in emerging fields. Altogether the Consortium members span across the whole technical and value chain required for developing and implementing solutions in terms of design, digitization, manufacturing, construction, and monitoring of resilient and sustainable buildings. The Consortium also includes partners with experience in social sciences, user engagement, and training to ensure the success and widespread application of new technologies in local communities. The Consortium will also support clustering activities with other relevant research projects to share knowledge and raise public awareness of building resilience. An international outreach and cooperation strategy will also be implemented to tackle the project challenges.

Table 1. Consortium

Number	Role	Short Name	Legal Name	Country
1	CO	TU Delft	TECHNISCHE UNIVERSITEIT DELFT	NL
2	BEN	PFE	PRIEDEMANN FASSADENBERATUNG GMBH	DE
3	BEN	IES R&D	IES R&D	IE
4	BEN	INCDFP	INSTITUTUL NATIONAL DE CERCETARE-DEZVOLTARE PENTRU FIZICA PAMANTULUI	RO
5	BEN	UNIROMA1	UNIVERSITA DEGLI STUDI DI ROMA LA SAPIENZA	IT
6	BEN	XLD	X-LAM DOLOMITI SRL	IT
7	BEN	STRESS	SVILUPPO TECNOLOGIE E RICERCA PER L'EDILIZIA SISMICAMENTE SICURA ED ECOSOSTENIBILE SCARL	IT
7.1	AE	UNINA	UNIVERSITA DEGLI STUDI DI NAPOLI FEDERICO II	IT
8	BEN	AMS Institute	STICHTING AMSTERDAM INSTITUTE FORADVANCED METROPOLITAN SOLUTIONS(AMS)	NL
9	BEN	PMB	MUNICIPIUL BUCURESTI	RO
10	BEN	ASM	ASM - CENTRUM BADAN I ANALIZ RYNKUSPOLKA Z OGRANICZONA ODPOWIEDZIALNOSCIA	PL
11	BEN	RoGBC	ASOCIATIA ROMANIA GREEN BUILDING COUNCIL	RO
12	BEN	RINA-C	RINA CONSULTING SPA	IT
13	BEN	UTBV	UNIVERSITATEA TRANSILVANIA DIN BRASOV	RO
14	BEN	ACER	AGENZIA CAMPANA PER L'EDILIZIA RESIDENZIALE	IT
15	BEN	Boom	BOOM BUILDS B.V.	NL
16	BEN	OMRT	OMRT BV	NL
17	BEN	ROTHO BLAAS SRL	ROTHO BLAAS SRL	IT
18	BEN	ARUP	ARUP BV	NL
19	BEN	Tecuci	MUNICIPIUL TECUCI	RO
20	BEN	Hölscher	DIPL.-ING. HPLSCHER GMBH & CO.KG	DE

2.2. Data Management Plan

This deliverable describes the data that will be used and generated within the MULTICARE project and the plan for their retrieval, storage, and management during and after the project. This plan will be updated periodically during the project development.

The Data Management Plan ensures that the data involved in the research is managed according to the best practices supporting Findable, Accessible, Interoperable, and Reusable (FAIR) principles. It allows the data to be a useful support to the demonstration of the MULTICARE solutions, to the uptake of them by third parties and further developments and to further research MULTICARE solutions beyond the scope of this project. Given that interoperability and scalability are explicit objectives in MULTICARE, prioritizing data and output management to adhere to FAIR principles becomes critical. The OGC and building SMART standards will be the starting point for the developments, as well as other open standards coming from other organizations, such as OMG and ISO, to enhance the compliance of solutions.

For project developments and internal collaboration, data are stored and exchanged within a repository in Microsoft Teams, which offers versioning and backup systems and serves as the main internal communication platform for the MULTICARE consortium. To ensure the production, documentation, and sharing of data and metadata adhere to open standards, relevant open standards will be prioritized. Additionally, comprehensive documentation and specifications will be provided and linked to the data and other outputs in both human and machine-readable formats whenever possible. Interoperability-supporting tools, such as APIs, will be used to enhance interoperability and reuse, even if the data or output cannot be shared entirely openly. Trusted and widely known repositories will be employed to share the data and other project outputs, with the source links further disseminated through the MultiCare project website and communication channels to improve findability and reuse.

This plan will be updated periodically during the project development and an annex reporting in detail the single datasets and outputs and their management and sharing details will be provided at the end of the project. This document is furthermore supported by the deliverable D4.3 which sets out the requirements for the Common Data Environment (CDE) as well as specific standards to ensure interoperability.

1.2.1. FAIR principles

The FAIR principles provide rules and criteria to improve (digital) data discoverability and reuse. A summary of the FAIR principles is herein reported, which are defined according to the following document: *Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.*

Findable

- F1. (Meta)data are assigned a globally unique and persistent identifier.
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

Accessible

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
 - A1.1 The protocol is open, free, and universally implementable
 - A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

Interoperable

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data

Reusable

- R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
 - R1.1. (Meta)data are released with a clear and accessible data usage license
 - R1.2. (Meta)data are associated with detailed provenance
 - R1.3. (Meta)data meet domain-relevant community standards

2. Data Summary

This section provides an overview of the datasets that the consortium plan to use and generate within MULTICARE. The following tables, providing the characteristics of each type of dataset, will be constantly revisited throughout the duration of the project.

2.1 Interviews and workshop data

Table 2 presents an overview of the data collected through interviews and workshops conducted as part of the MULTICARE project. These data types collectively provide insights into various aspects of indoor environmental quality (IEQ), user preferences and satisfaction, facilitating informed design and decision-making for the demonstrator buildings.

Table 2. Interviews and workshop data

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
Measurement and verification protocols	Spreadsheets	CSV	Assessing IEQ under different building controls and understand impact on the operation of the buildings from users	<5gb	Collected during the project	Re-use for further related investigations
Metadata on how the data has been collected and information on sensing technologies	Metadata	TXT	Ensure FAIR data sharing	<5kb	Generated during the project	Re-use for further related investigations
Measurement of subjective user preferences, expectations on multi-hazard resilience, IEQ, satisfaction with design	Questionnaires	CSV	Evaluate the satisfaction of users, perception, preferences in the space with IEQ, OEQ, resilience, satisfaction with proposed design	<5kb	Collected during the project	Current use
User data regarding age and gender and	Questionnaires	CSV	Evaluate the link between gender, age	<5kb	Collected during the project	Re-use for further related investigations

country of provenience of the end users in demo buildings and sites			and country of provenanced and satisfaction of users in the space with IEQ			
List of volunteers, name, and emails, with information on whether the consent form has been signed or no	Spreadsheets	CSV	Ask for Consent and feedback	<5kb	Collected during the project	Current use
Transcripts from workshops and interview, already anonymised	Spreadsheets	CSV	Ask for Consent and feedback	<5kb	Collected during the project	Current use

2.2 Existing data

Table 3 presents an overview of the existing datasets purposed for the MULTICARE project. Collectively utilized in the carrying out of multi-criteria risk and resilience assessment of the as built building or urban scenario.

Table 3. Existing Data

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
Hazard maps	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB	Compilation of existing hazards maps	<50mb	Literature, reports	Risk assessments
Energy urban data	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	Assessment of urban energy performance	<50mb	Literature, reports	Energy performance and assessments
Exposure models	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	Assessment of urban resilience	<100mb	Literature, reports, census data	Risk assessments

Socio-economic urban data	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	Assessment of urban resilience	<50mb	Literature, reports, census data	Various
Hazard data	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, CSV, EPW	Compilation of existing hazards data	<50mb	Literature, reports	Risk assessments
Fragility data	Tabular	CSV, XLSX, XML, JSON, SQL	Compilation of existing fragility data	<50mb	Literature, reports	Risk assessments
3D geospatial data	Geospatial	3D GML, CityGML, LAS	Base data for risk assessments, exposure models and early warning systems	<10gb	Existing data augmented by data collected during the project	Research, analysis and dissemination

2.3 Generated, synthetic and sample data

Table 4 presents an overview of the data generated by the MultiCare project. Such data includes the content of deliverables, results of simulations and other processes. Data may be generated as in the results of say, a model output; or be the result of research and development of methodologies.

Table 4. Generated, synthetic and sample data

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
KPIs tables	Spreadsheets	XLSX	Definition of KPIs for technological and digital solutions	<0.5mb	Literature, surveys	Performance indicators for holistic assessment
Measurement and verification protocol	Documents	PDF	Evaluation framework for baseline vs retrofitted buildings	<2mb	Literature	Energy savings assessment
Resilience table	Spreadsheets	XLSX	Definition of resilience indicators	<1mb	Literature, surveys	Indicators for resilience assessment

Energy building data	Tabular	CSV, XML, JSON	Assessment of building energy performance	<50mb	Literature, reports	Energy performance and assessments
Socio-economic building data	Tabular	CSV, XML, JSON	Assessment of socio-economic building losses	<50mb	Developed during project	Various
Resilience building data	Tabular	CSV, XML, JSON	Assessment of building resilience	<50mb	Developed during project	Various
Multi-hazard resilience quantification	Code / scripts / models	PY, BAT, JSON	Assessment of building resilience	<50mb	Developed during project	Resilience assessments
Plug-ins for multi-criteria assessment	Code / scripts / models	PY, GHX	Assessment of building resilience	<50mb	Developed during project	Resilience assessments
Building models	Code / scripts / models	IDF, TCL, 3DM, OBJ, DWG	Modelling of case studies	<1gb	Developed during project	Various
Database of façade materials, connections, and systems	Databases	SQL	Resilience based façade design	<500mb	Developed during project	Façade design, risk assessments
Calibrated vulnerability models	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	Risk assessments	<200mb	Improvements on existing data	Research, analysis and dissemination
Calibrated exposure models	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	Risk assessments	<200mb	Improvements on existing data	Research, analysis and dissemination
Energy simulation results	Tabular	CSV, XLS	Results from energy simulations	<100mb	Developed during project	Research replicability and results reuse
Results of numerical simulations	Tabular	CSV, XLS	Results from numerical simulations	<100mb	Developed during project	Research replicability and results reuse

Finite Element Analysis (FEM) models – connections & sub-assemblies	Code / scripts / models	INP, THM, TXT	Models for numerical simulations	<5gb	Developed during project	Research replicability and results reuse
Models for as built and intervention	BIM Models	IDF, TCL, 3DM, OBJ, DWG	Modelling of case studies	<1gb	Developed during project	Various
Models for as built and intervention	BIM Models	IDF, TCL, 3DM, OBJ, DWG	Modelling of case studies	<1gb	Developed during project	Various
Technical reports	Documents	PDF	Report on as-built performance	<2mb	Municipality data	Current use
Technical reports	Documents	PDF	Report on as-built performance	<2mb	Municipality data	Current use
Technical drawings	Images	JPEG, PNG, DWG	Structural details of buildings	<50mb	Municipality data	Current use
Technical drawings	Images	JPEG, PNG, DWG	Structural details of buildings	<50mb	Municipality data	Current use
Urban heat vulnerability models	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	Results from analysis, used for virtual intervention planning	<1gb	Developed during project	Research replicability and results reuse
Urban heat intervention models	Tabular	CSV, XLS	Investigating impact of proposed retrofitting options.	<200mb	Developed during project	Research replicability and results reuse
SDSS geospatial results	Geospatial	GeoJson, GDB, Shapefile, GeoTIFF, GML, WKT, WKB, XLSX	The geospatially encoded data describing the potential impact of the solutions at an urban	<500mb	Developed during project	Research replicability and results reuse, dissemination

			scale.			
SDSS tabular results	Tabular	CSV, XLX	Tabular results describing the potential impact of the solutions at an urban scale.	<500mb	Developed during project	Research replicability and results reuse, dissemination

2.4 Collected data

Table 5 presents data collected by the MultiCare project. This data pertains geospatial and building information data collected via sensors, drones, and other such tools.

Table 5. Existing Data

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
Sensor time-series data	Spreadsheets	CSV, JSON, XML	Various time-series datasets used in risk assessments, predictive warning systems and performance analysis.	<500mb	Collected during the project	Research and analysis
Sensor metadata	Metadata	TXT, RTF, JSON, XML, HDF5	Analysis of sensor results	<10mb	Collected during the project	Research and analysis
Drone geospatial data	Geospatial	GeoTIFF, LAS	Base data for risk assessments, exposure models and early warning systems.	<100gb	Existing data augmented by data collected during the project	Research, analysis and dissemination

2.5 Digital tools

Table 6 presents a series of isolated digital tools developed by the MultiCare project. This includes web-services and algorithms.

Table 6. Digital Tools

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
Sensor data webservice	Code / scripts	Digital repository (multiple scripting languages)	Cloud based web-service for pushing / pulling sensor time series data	<200mb	Developed within the project	Continued collection of sensor data post-project (or integration with future sensor setups)
Overheating predictive algorithms	Code / scripts	Python scripts	Predictive algorithms for overheating risk	<10mb	Developed within the project	Research, analysis and dissemination
Travel time estimator	Code / scripts	Digital repository (multiple scripting languages), API	As part of deliverable D10.5	<1gb	Developed within the project	Dissemination and use by stakeholders

2.6 Documentation, Guides and Educational Material

The data and materials are intended to underpin a shared multidisciplinary knowledge base aimed at developing standardized frameworks and effective solutions for enhancing the resilience of buildings. Furthermore, they will support the training and educational initiatives developed and delivered within the MultiCare project (**Table 7**. Documentation, Guides and Educational Material).

Table 7. Documentation, Guides and Educational Material

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
Measurement and verification protocols	Documents	PDF	Evaluation framework for baseline vs retrofitted buildings	<2mb	Literature	Energy savings assessment
Guidelines on façade and exoskeleton solutions	Documents	PDF	Documentation on technology detailing and assembly	<5gb	Developed during project	Uptake of façade and exoskeleton solutions
Handbook for resilient structures	Documents	PDF	Documentation on designing low-damage structural connections	<5gb	Developed during project	Uptake of low-damage structural systems
Lab reports	Documents	PDF	Results from experiments	<5gb	Developed during	Research replicability

					project	and results reuse
Technical reports	Documents	PDF	Report on as-built performance	<2mb	Municipality data	Current use
Training materials	Documents	PDF	Capacity building	<5gb	Developed during project	Knowledge on resilience design, assessment, management
	Images	JPEG, PNG	Capacity building	<5gb	Developed during project	Knowledge on resilience design, assessment, management
	Digital Video	MP4	Capacity building	<100gb	Developed during project	Knowledge on resilience design, assessment, management

2.7 Communication and Dissemination

These data are generated, utilized, and shared for the purpose of communication and dissemination and are listed in **Table 8**.

Table 8. Communication and Dissemination

Data description	Type	Format/s	Utility within Project	Expected Size	Source	Utility outside the project
Papers	Documents	PDF	Communication and dissemination	<1gb	Developed during project	Research replicability and results reuse
Presentations	Documents	PPT, PDF	Communication and dissemination	<5gb	Developed during project	Research replicability and results reuse
Website	Documents	HTML	Communication and dissemination	<100gb	Developed during project	Support project results uptake
Brochures	Documents	HTML	Communication and dissemination	<100gb	Developed during project	Support project results uptake
Booklet	Documents	HTML	Communicate about low-carbon resilient solutions	<100gb	Developed during project	Support project results uptake

3. FAIR data

3.1 Making data Findable, including provisions for metadata

All necessary measures will be implemented to ensure the findability of datasets produced within the MULTICARE project. To achieve this goal, each generated dataset will be assigned a globally unique and persistent identifier. Additionally, rich and standard-compliant metadata, including indexed search keywords facilitating discovery, will be attached to the datasets. In order to comply to such conditions, every dataset will be assigned a Digital Object Identifier (DOI), to make them citable and persistently available.

The data will be published through trusted repositories such as Zenodo³ or 4TU.ResearchData or Open Research Europe. Such repositories are based in Europe and are certified with a Data Seal of Approval certification. They can assign a globally unique and persistent identifier in form of a DOI. In addition, they allow integration with GitHub repositories, so that open code that will be possibly further developed after the project could be published and referenced as well. All the published datasets will be accompanied by rich standardised and community-endorsed metadata⁵ (adhering to DataCite, Dublin Core or other relevant and standard-compliant metadata schema such as DCAT, GeoDCAT, ISO19115 for spatial datasets) to ensure that all datasets are findable. In addition, to further aid their discoverability, keywords describing the datasets will be added. 4TU.ResearchData is also using schema.org metadata, meaning that all datasets are indexed in Google Dataset Search.

In case of papers, publication venues will be chosen such that a persistent identifier can be assigned to the documents and indexing in relevant databases (e.g. Scopus, WoS) will help find the resources. The connection and advertisement of the publications to relevant repositories and research exchange platforms, such as Google Scholar, ResearchGate, and social media, such as Twitter or LinkedIn will also improve and support the resources findability. Other media data will be published through other kinds of repositories and sharing platforms, such as YouTube for videos through the MULTICARE YouTube channel. Such platforms provide a URL and identifier which can be used to identify the resources uniquely. They also provide maintenance and ensure persistence of the published resources. The MULTICARE website will be used to link all the generated datasets and possibly publish further resources which do not belong to the previous categories.

The developed software in MULTICARE will be published as Free Open Source (40%) or commercial (60%). The free and open-source software will be developed in GitHub and published through Zenodo or 4TU.Research, getting assigned a DOI. Commercial software websites will be linked to the MULTICARE website together with related documentation. The standard documents (standards or best practices) developed within the project will be published through the standard development organisations repositories and a persistent identifier will be assigned by their resource management systems (e.g. OGC Rainbow⁶).

3.2 Making data accessible

The publication of data through the trusted repositories (e.g. Zenodo, 4TU.Research) will also support data accessibility, mainly through the web. They use HTTPS protocol (Hypertext Transfer Protocol Secure) which is based on TCP/IP. All the data generated within the project will be made openly available and accessible through the publication repositories, except for part of the software code, according to the involved companies' business models and exploitation plans. Metadata for all the data, containing the information to access the data, will be made openly available and licensed under a public domain dedication CC0, as per Grant Agreement. In 4TU.ResearchData, metadata may be freely reused under the CC0 waiver. Information about the accessibility of the dataset is also included in the metadata. Datasets in 4TU.ResearchData will be openly available and can be accessed directly from the dataset's landing page.

The data published in 4TU.ResearchData, will be archived and preserved within the repository for at least 15 years. The accessibility of the developed software in the project will vary according to the business models of the developing companies. It will be provided as: free open source (30%) – open access; open source (30%) - access after payment; proprietary (40%).

3.3 Making data interoperable

The use and development of standards will be a key part of the project and will be essential to allow interoperability and enable scalability of results. In particular, the standards developed by the Open Geospatial Consortium (OGC) and building SMART will be considered. A core part of the project development regards the use and development of standard data models (e.g. OGC CityGML, CityJSON, 3D-Tiles, GeoTiff, bSI Industry Foundation Classes, EXPRESS, gltf, json), best practices and interoperability supporting tools (e.g. OGC and bSI APIs, OGC Rainbow, bSI Data Dictionary, Information Delivery Specification, Information Delivery Manual, mvdXML, etc.).

Other ISO standards, such as ISO TC211, ISO 19650 and INSPIRE prescriptions could also be used as an additional reference for some of the developments. All the standards profiles and specifications will be openly made available through publication in open repositories or through the dedicated channels in the standard development organizations (e.g. OGC and buildingSMART) to allow reuse and further improvement by the community and by further future studies. The data and related metadata will include references (as DOI or another identifier, such as URL) to other data when relevant to understand the data provenance and context. The developed software will strain towards interoperability by reusing, where possible, the available libraries and resources, using open data protocols, format and web interfaces. It will be developed either as Open Source software (50%), making the source code available for any extension and further improvement, or as a proprietary software modules / services. In the latter case (50%), they will connect to the central management platform through APIs published through standard OpenAPI specifications, which will contribute to making the overall architecture interoperable.

3.4 Increase data re-use

Whenever possible, all the data will be shared in their editable format, so that they can be easily re-used, modified and extended for further studies and applications and to support scalability of solutions. Metadata about the data and any further documentation or specification useful to access, understand and re-use the data, including README files, will be published as part of the same resource in the data repository. In addition, the methodologies and processes used to obtain the resulting (published) datasets (provenance or lineage) will be explained in either scientific publications or internal reports to be attached to the datasets themselves, allowing replicability and reuse.

All the data, possibly anonymized, in case of interview or workshop data, will be licensed using standard reuse licences allowing free and open access and re-use (e.g. CC0, CC BY) and will remain openly usable after the end of the project. In case of software, licences will be attached according to the software companies business models. It will result approximately in: free open code license + docs (60%); Code sharing after the payment of a fee + docs (30%); no code sharing + docs (10%). One possible standard software license that will be used for proprietary software will be MIT licence, while open source software will mainly be provided under MIT and GPL-3.0 licenses.

To ensure data quality, the data, as well as software, will be validated using suitable methodologies according to the specific kinds of data (e.g. workshop-based human validation, data models validation through standard schema validators, iterative tests, iterative feedback collection, internal review and feedback processes and so on).

4. Other Data Management Provisions

4.1 Allocation of resources

TU Delft will be the main responsible party for data management storage and quality assurance for the project. A part of the budget is reserved for the development of the data management and storage platform during the project. The chosen platforms for publishing the final data are available freely available. 4TU.ResearchData repository ensures data quality and curation (manual curation at the time of deposition, and automated curation and checks for data integrity after the deposit) and can archive 1TB of data per researcher per year free of charge for all TU Delft researchers. We do not expect to exceed this and therefore there are no additional costs of long-term preservation. GitHub also allows free repository and maintenance for the hosted resources, as well as the other potential repositories chosen. Some specific standard related data and products could be stored alternatively in the OGC or building SMART repositories. This would be part of current standard development organisations' business and resources; therefore, no additional resources are required from the project.

4.2 Data security

The chosen repositories to store and publish the data have internal security measures which allow safe storage, ensuring the transfer of responsibility if needed as well as backup functionality. They are also enabled to manage accesses. No highly confidential data requiring encryption will be managed within the project.

4.3 Ethics

The ethics related to the data collected and generated in the MULTICARE project are addressed in deliverable D1.3.

5. Conclusion

This DMP has presented a comprehensive list of the expected data categories that will be utilized or produced during the project, as well as their expected type, size, format, and utility within and without the project. Furthermore, the guidelines for maintaining conformity to the FAIR data principles have been set out. This document is further supported by the deliverable D4.3 (Common Data Environment, Interoperability and Standards) which lays out the specific technical standards to ensure interoperability.

The DMP shall remain a live document, undergoing revisions and updates in parallel with project developments.