PROBONO

D10.11 Data Management Plan and GDPR Compliance (I)

Deliverable nº:	D10.11
Deliverable name:	Data Management Plan and GDPR Compliance (I)
Version:	1.0
Release date:	**30/09/2022**
Dissemination level:	Public
Status:	Peer-reviewed (Draft, Peer-reviewed, Submitted, Approved)
Author:	**AKKA – Samantha DRAY**

Project abstract:

This document defines all the procedures to handle the data collected or generated and how they are processed and preserved in the PROBONO project.

Document history:

Version	Date of issue	Content and changes	Edited by
0.1	02.03.2022	Initial Deliverable Structure	Samantha DRAY (AKKA)
0.2	05.03.2022	Draft of sections 1-3 and 5	Samantha DRAY (AKKA)
0.3	15.04.2022	Update to deliverable draft	Samantha DRAY (AKKA)
0.4	10.05.2022	WP3, WP5, WP6, WP7, WP9 contributions	Samantha DRAY (AKKA)
0.5	20.07.2022	WP1, WP2, WP4, WP8 contributions	Samantha DRAY (AKKA)
0.6	31.07.2022	Final updates	Samantha DRAY (AKKA)
0.7	26.08.2022	Change of template for the document	Samantha DRAY (AKKA)
0.8	28.09.2022	Review	Ivan Fratrić (CGBC) & Javier Dorao (DCN)
0.9	30.09.2022	Rework	Samantha DRAY (AKKA)
1.0	30.09.2022	Submission	

Peer reviewed by:

Partner	Reviewer
CGBC	Ivan Fratrić
DCN	Javier Dorao

Deliverable beneficiaries:

WP / Task
ALL

DEFINITIONS[1]

A Green Building (GB) (new or retrofit) is a building that, in its design, construction and operation, reduces or eliminates negative impacts, and can create positive impacts, on the climate, social, and natural environment. GBs preserve precious natural resources and improve quality of life[1]. Specifically, this means that GBs should be very energy efficient, use extensively the potential of locally available renewable energy, use sustainable materials, and aim for a low environmental impact over the entire life cycle. GBs offer their users and residents a healthy climate and a high quality of stay, they are resilient e.g., to environmental change and contribute to social inclusion.

Green Neighbourhoods aligned with the European Green Deal[2], is a set of buildings over a delimited area, at a scale that is smaller than a district, with potential synergies, in particular in the area of energy. A green neighbourhood is a neighbourhood that allows for environmentally friendly, sustainable patterns and behaviours to flourish e.g., bioclimatic architecture, renewable energy, soft and zero-emission mobility etc. Green neighbourhoods are the building blocks of Positive Energy Districts (PEDs)[3] by implementing key elements of PED energy systems. For example, the exchange of energy between buildings increases the share of local self-supply with climate-neutral energy and system efficiency. They also provide the technical conditions to enable Citizen Energy Communities[4] and Renewable Energy Communities[5] to be implemented

Green Buildings and Neighbourhoods (GBN) in PROBONO are GBs integrated at delimited area or district level with green energy and green mobility management and appropriate infrastructure supported by policies, investments and stakeholders' engagement and behaviours that ensures just transition that maximise the economic and social cobenefits[1] considering a district profile (population size, socio-economic structure, and geographical and climate characteristics). Delivered in the right way, GBN infrastructure is a key enabler of inclusive growth, can improve the accessibility of housing and amenities, reduce poverty and inequality, widen access to jobs and education, make communities more resilient to climate change, and promote public health and wellbeing.

DGNB certification serves as a quality stamp ensuring the state of the building for buyers. The Green Building Council Denmark[6] has established the German certification DGNB meaning 'German Society for Sustainable Buildings'. The Danish version of DGNB was created to obtain a common definition of what sustainability is towards and making it measurable. A consortium of experts was established from all parts of the construction sector. DGNB had to be reshaped for the Danish standards, practice, traditions, and laws but is now available to certify any construction project. They chose DGNB as an innovation-forward and sustainable future guarantee. DGNB diversifies itself by focusing on sustainability and not just the environment. DGNB creates a standardised framework for the construction operations conditions and creates a common language which facilitates communication between professions and helps organize and prioritize the efforts in long and complicated development phases.

Life cycle assessment (LCA)[7] is a tool used for the systematic quantitative assessment of each material used, energy flows and environmental impacts of products or processes. LCA assesses various aspects associated with development of a product and its potential impact throughout a product's life (i.e. cradle to grave) from raw material acquisition, processing, manufacturing, use and finally its disposal. In PROBONO, LCA represents the statement of a building's total energy, resource consumption and environmental impact in the manufacture, transport, and replacement of materials and for its operation over its expected life. Social life cycle assessment (S-LCA)[8] is a method to assess the social and sociological aspects of products, their actual and potential positive as well as negative impacts along the life cycle. Life-cycle costing (LCC)[9] considers all the costs incurred during the lifetime of the product, work, or service.

---

[1] These definitions are identified in the PROBONO GA no 101037075 and they may be further refined according to the results of the project.

Table of contents

Abbreviations and Acronyms

Acronym	Description
BIM	Building Information Modelling
CA	Consortium Agreement
CC	Creative Commons
D	Deliverable
DGNB	German Society for Sustainable Buildings
DMP	Data Management Plan
DT	Digital Twin
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable and Reusable
GB	Green Building
GBN	Green Building Neighbourhood
HVAC	Heating-Ventilation-Air Conditioning
ISO	International Standards Organization
KPI	Key Performance Indicators
LCA	Life Cycle Assessment
LL	Living Lab
OSI	Open Source Initiative
ORDP	Open Research Data Pilot
PV	Photovoltaics
RTD	Research and Technical Development
TRL	Technology Readiness Level
V	Version
WP	Work Package

Project Partners

Organisation	Abbreviation	Country
INGENIERIA ESPECIALIZADA OBRA CIVIL E INDUSTRIAL SA	ACC	Spain

INLECOM INNOVATION ASTIKI MI KERDOSKOPIKI ETAIREIA	INLE	Greece
FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.	FRHF	Germany
SERCO BELGIUM	SERCO	Belgium
MOTT MACDONALD FRANCE	MM	France
AARHUS UNIVERSITE	AU	Denmark
AKKA HIGH TECH	AKKA	France
INSTITUTO TECNOLOGICO DE ARAGON	ITA	Spain
GECO GLOBAL APS	GECO	Denmark
UNIVERSITY COLLEGE DUBLIN, NATIONAL UNIVERSITY OF IRELAND	DUBLIN UCD	Ireland
DUN LAOGHAIRE RATHDOWN COUNTY COUNCIL Ireland	DLR	Ireland
INSTITUT DE RECHERCHE TECHNOLOGIQUE SYSTEMX	IRTSX	France
SIXENSE ENGINEERING	RESAL	France
KONNECTA SYSTEMS LIMITED	KNT	Ireland
VLTN GCV	VLTN	Belgium
EBOS TECHNOLOGIES LIMITED	eBOS	Cyprus
POLYTECHNEIO KRITIS	TUC	Greece
TPF UTILITIES	TPF	Belgium
CONSEIL INTERNATIONAL BIODIVERSITE & IMMOBILIER	IBPC	France
SMART INNOVATION NORWAY AS	SIN	Norway
DIN DEUTSCHES INSTITUT FUER NORMUNG E.V.	DIN	Germany
FUNDACION CARTIF	CARTIFF	Spain
PNO INNOVATION SL	PNO	Spain
ERASMUS UNIVERSITEIT ROTTERDAM	EUR	Netherlands
VIAS INSTITUTE	VIAS	Belgium
UNIVERSIDAD DE SANTIAGO DE COMPOSTELA	USC	Spain

BOVLABS SAS	BLABS	France
STAM SRL	STAM	Italy
DISTRITO CASTELLANA NORTE SA	DCN	Spain
AYUNTAMIENTO DE MADRID	MAD	Spain
SOPREMA	SOP	France
COWI AS	COWI	Denmark
DE L'AUTRE COTE DE L'ECOLE	ACE	Belgium
VISBLUE APS	VISB	Denmark
COMPANIA ESPANOLA DE LAMINACION SL	CELSA	Spain
FUNDACION CIDAUT	CIDAUT	Spain
TELCOSERV SCHEDIASMOS YLOPOIHSH & BELTIOPOIHSH THLEPIKOINONIAKON DIKTIONEGKATASTASIS THLEPIKOINONIAKOU EKSOPLISMOU DIATAKSEONKATASKEVASTIK	TSRV	Greece
IDOM CONSULTING, ENGINEERING, ARCHITECTURE S.A.U.	IDOM	Spain
SONAE MC - SERVICOS PARTILHADOS, SA	SONAE	Portugal
WILLIS TOWERS WATSON SA	WT	Belgium
ECOFOREST GEOTERMIA SL	ECOF	Spain
ANERDGY AG	ANERDGY	Switzerland
BEEPLANET FACTORY SL	BEE	Spain
CAPWATT, S.A.	CAPW	Portugal
HRVATSKI SAVJET ZA ZELENU GRADNJU	CGBC	Croatia
CESKE VYSOKE UCENI TECHNICKE V PRAZE	CTU	Czech Republic
MESTSKA CAST PRAHA 6 / District Prague 6	PRAGUE	Czech Republic

Executive summary

This deliverable report, D10.11 Data Management Plan and GDPR Compliance (I) is an output of the Project Management Work Package (WP10, Task 10.4). It guides the PROBONO project beneficiaries with regard to data collection, protection, minimisation, anonymisation, and codification of recognisable information, storage, transfer and analysis of digital and physical personal data subject to European and national legislations and GDPR.

The report includes a thorough description of the data that will be generated, processed or used within PROBONO and details on the category and nature of the data resulting from each of the PROBONO work packages. The methodology applied within PROBONO ensures that the project will follow 'FAIR' data principles as defined by the EC and these are also included in this report.

The data lifecycle steps are explained in the report, including the approach to Data collection, processing, Use and Storage. The roles and responsibilities are given as follows:

- **Ethics and Data Protection Officer (EDPO)** - Responsibility for leading the data management related activities, reviewing the Data Management Plan (DMP) and policies and providing guidance as necessary.

- **Data Controller** - Appointed by each Work Package generating or using data within the project. Responsible for ensuring that the DMP is implemented within their organisations.

Tables of the data that are expected to be collected and used are provided per partner per work package. A chapter of the DMP focuses on GDPR and the protection of personal data. This section includes all policies that PROBONO will follow to comply with the above.

Annex I provides a data management report which will be completed and included in the subsequent versions of the DMP report.

The Data Management Plan will be stored in Teamwork under WP10's folder. An updated version of the deliverable will be produced at M30 before the mid-term of the project with a final version being produced at M60.

1   Introduction

1.1   Project Overview

PROBONO brings together a European multidisciplinary consortium of 47 partners, construction and consulting entities, public asset service managers, municipalities, technology solution providers and experts, with the goal of turning six European districts (PROBONO´s Living Labs) into Green Building Neighbourhoods (GBN), which will have a positive energy balance and zero carbon emissions. They are divided into two large-scale demonstrators (Madrid and Dublin) and four living labs representing business/owner promoters of the green buildings and neighbourhoods' transition (Porto, Brussels, Aarhus, Prague).

PROBONO will provide strong examples of how GBN's technological and social innovations can be applied, along with a vision focused on building infrastructure, as well as a renewed focus on people and sustainability, while taking full advantage of digitalization and smart technologies for the benefit of society. The adoption of the PROBONO approach and innovations will be proposed through a range of participatory methods that promote stakeholders (including citizens) partaking in co-designing and co-delivering a sustainable GBN. PROBONO will provide GBN Strategic Planning Tools in spatial, economic, technical, environmental regulatory, and social context, all of which will be aligned with city and urban masterplans and policy frameworks. The project will create evidence-based policy recommendations, standardization actions, and robust adoption and commercialization strategies supported by a capacity-building program and a European Alliance of GBN Innovation Clusters.

PROBONO will provide a GBN Digital Twin (DT) implemented across all the LLs and will serve as a virtual representation of associated GBN's, including operational assets that implicate environmental and efficiency KPI. A cloud-based decision support-planning tool will be created to develop an optimized design for carbon-neutral energy GBN systems incorporating PROBONO's innovative solutions on GBN demand and response dynamics. The technological developments will include metering different utilities such as electricity, gas, warm energy, cold energy, and water linked to Smart IoT gateway and Energy Optimisation middleware, all this combined with geothermal, PV, micro-turbines, efficient HVAC technologies, green roofs, custom insulation, GB energy optimization, efficient energy storage, and integrating an EV charging value chain. To enhance wide-scale adoption and standard creation, PROBONO will provide evidence-based policy recommendations, standardization actions, and robust adoption and commercialization strategies.

1.2   Mapping PROBONO Outputs

Purpose of this section is to map PROBONO's GA commitments, both within the formal Deliverable and Task description, against the project's respective outputs and work performed.

Table 1: Adherence to PROBONO's GA Deliverable & Tasks Descriptions

GA Component Title	GA Component Outline	Respective Document Chapter(s)	Justification
TASKS			

	Development of the PROBONO DM plan, which will include information related to the types of data and metadata the project will generate and collect, the standards that will be used to represent the data during the project and how partners might exploit the data resulting from the project in alignment with the Innovation management and IPR protection procedures, GDPR Policies, the EC FAIR principles and the guidelines of the Ethics, Legal and Security Office.		Details and explanations on the implementation of the data management plan in PROBONO.
Task 10.4 - Data Management, GDPR and ethics respecting policies	The DM manager will act as the PROBONO Data Protection Officer (DPO). The DPO will set management procedures to protect personal data collected as a result of project activities, as well as buildings critical data, from unauthorised use or sharing.	Chapter 2 to 6	Data handled by the project and every work packages. Methodologies and standards used in the project. Access and curation of the data handled.
	The DPO will manage the ethical issues, and tasks as the design of the information and consent forms if requested (e.g., a research experiment involves humans), and will deliver ethical and legal guidelines early in the project to stipulate the principles that PROBONO will adhere to.		
DELIVERABLE			
<u>D10.11: Data Management Plan and GDPR Compliance (I)</u> This report formulates the findings of T10.4, and explores the Initial data management plan, definition of data and characterization, procedures & related policies in respect to the GDPR.			

1.3   Purpose of the deliverable

The Data Management Plan defines all the procedures to process and preserve all collected or generated data,. It describes the approach to making the PROBONO data ''FAIR'' by indicating what kind of data will be generated, collected, and processed. Also, it determines what standards will be applied, how research data will be preserved, and what parts of the datasets will be shared for evaluation purposes and compliance with ORDP requirements. The document addresses ethical and confidentiality matters, as well as some data security principles. This deliverable is a living document and will be updated as the project evolves:

- The initial version D10.11 (M6, i.e. July 2022) describes the data management plan according to the current evolution of the project. It identifies an initial set of data categories that will be involved in the project and proposes the data management process that will be followed in future developments.

- The updated version D10.12 (M30, i.e. July 2024) will provide an update of the data management plan, including the description of the datasets and the possible evolution of the procedures defined early in the project.

- The final version D10.13 (M60, i.e. January 2027) will include the final description of the datasets and procedures.

1.4   Intended audience

The dissemination level of this D10.11 deliverable is 'public' (PU). AKKA as WP10/Task 10.4 leader is responsible for it, and its main contributor, with the assistance of all partners of the consortium.

Appointed peer reviewers are DCN and ANE.

1.5   Structure of the deliverable and its relation with other work packages/deliverables

This deliverable serves as an entry point to understanding the project-wide approach to data management in the PROBONO project. It provides an overview of data management at the project level, with a stronger focus on research data management as is required by the H2020 ORDP programme. Later versions will also cover categories of GBD technology, construction, and energy data that will be used in the course of the project.
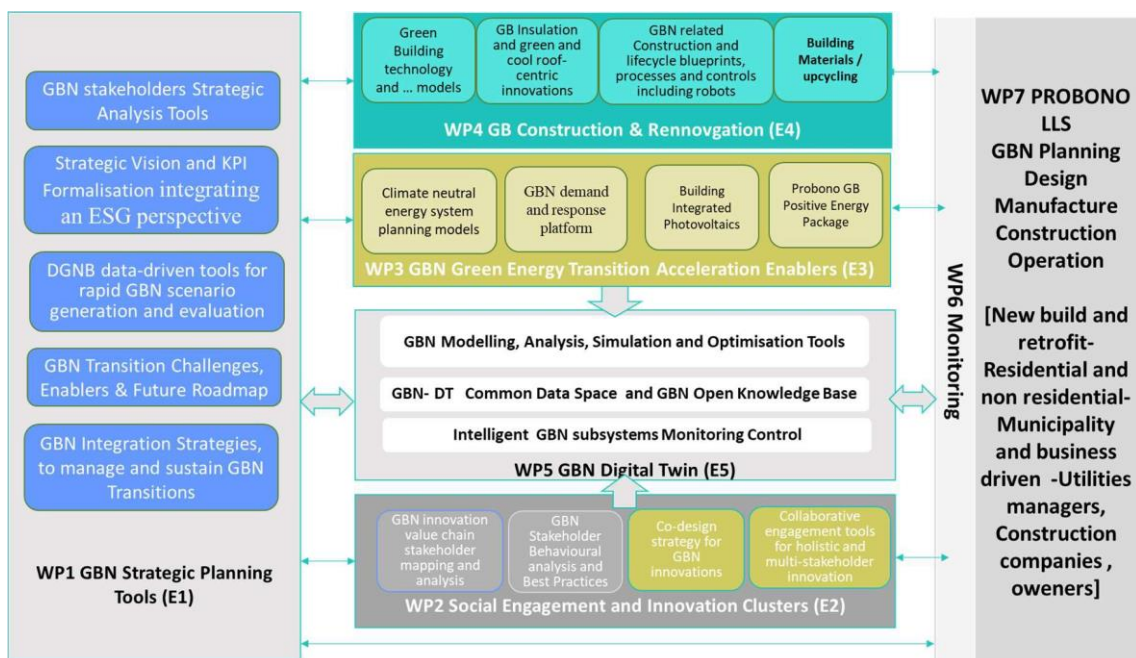
Figure 1: PROBONO Conceptual Architecture and relation to the Workplan (WPs)

1.6   PROBONO data management with DMPOnline

DMPOnline is a web tool that facilitates the collaboration and sharing of the DMP. Access can be given to people using different profiles: reader, writer or owner. The site also permits the creation of multiple revisions of the document.

This plateform has been chosen for PROBONO to provide an easy and efficient way for partners and Work Package leaders to share their inputs and review them and thereby having a living document.

Figure 2 - DMPOnline for PROBONO

2 Data Handled by the Project

PROBONO will handle different types of data which can be organised into four categories:

- administrative data

- open research data

- evaluation data

- technical data

This chapter will describe these data categories while providing information on their management and sharing.

2.1 Administrative data

This category refers to the data produced by the project management activities within WP10, such as meeting minutes, recordings, internal reports (for historical purposes) and follow-ups. In terms of deliverables, they include:

D10.1 Project Management Handbook and Yearly Management Reports

D1.2 and D1.3 Project - Innovation and Quality management plan.

The data is collected by the management team, which includes the project manager, the WP leaders, and task leaders. It is stored in Microsoft Teams, which is a collaborative application that provides different basic functionalities and extensions. Preliminary guest accounts have been requested for all the members of the Consortium so they can have access to all available information linked to the Project. This administrative data is confidential and is available only to the members of the Consortium.

On the other hand, the administrative data refer also to another type of data linked to project management activities, including:

- Data management: D10.11, D10.12 and D10.13 Data Management Plan
- Project Management: D10.1 Project Management Handbook and Yearly Managements reports
- Project activities communication: WP8
- Activities concerning project stakeholders: WP1

This data is open to the public and is accessible via the project website.

2.2   Open Research data

The open research data category refers to the data collected from the research work, as well as WP7. It includes D5.11 PROBONO Open Data Sources from LLs. These data are open to the public and are accessible via the repository chosen by the Consortium.

2.3   Evaluation data

The evaluation data category includes technical data concerning the quantitative and qualitative evaluations and other related business models for data sharing services. They include KPIs measured for the data sharing culture in order to determine a list of requirements that promote data sharing in the context of exploration and collaboration among actors, as well as a projection of the data sharing ecosystem in GBN's for the coming years in terms of qualitative estimations.

The evaluation data will be accessible to the public via a repository of the Consortium's choice.

2.4   Technical data

The technical data category includes data related to the technical developments of the PROBONO project, which corresponds to the contents of WP5.

In WP3, the project will evolve towards actual and real-time data usage that will be fostered by physical Living Labs' implementations in different locations, and a Virtual Lab that will support the process and foster data exchange for the entire duration of the project.

The technical data mentioned above can fall into two categories: *Public* or *Confidential* (only for members of the consortium, including the Commission Services).

The following table is a template that can be used to describe the datasets of the project.

Dataset Reference	**Each dataset will have a reference that will be generated by the combination of the name of the project, the trial site, the use case in which it is generated and the datatype: "PROBONO_Trial-Site_UC_Datatype".**
Dataset Name	**Name of the dataset.**
Dataset Description	**Each dataset will have a full data description explaining the data provenance, origin and usefulness. Reference may be made to existing data that could be reused.**

Standards and metadata	**The metadata attributes list to be used to find the dataset.**
File format	**All the format that defines data.**
Data Sharing	**Explanation of the sharing policies related to the dataset between the next options:** **Open: Open for public disposal.** **Embargo: It will become public when the embargo period applied by the publisher is over. In case it is categorized as embargo the end date of the embargo period must be written in DD/MM/YYYY format.** **Restricted: Only for project internal use.** **Each dataset must have its distribution license.** **Provide information about personal data and mention if the data is anonymized or not. Tell if the dataset entails personal data and how this issue is taken into account.**
Archiving and Preservation	**The preservation guarantees and the data storage during and after the project (for example databases, institutional repositories, public repositories, etc.).**

TABLE 2: DATASET DESCRIPTION TEMPLATE

3   Data Handled by Work Package

As the project is still in its early stage, we do not have a clear vision of the data that will be handled. As the DMP is a living document, we have made the choice to give hypotheses concerning the data handled. However, it is important to take that information with precaution as they are still hypotheses to be verified in the D10.12 DMP V2 (M30) and the D10.13 DMP V3 (M60).

3.1   WP1 & WP2

WP1 has the following five objectives:

- Provide value-oriented analysis tools and decision support system enabling GBN Stakeholders to:
  - Specify value-adding sustainability perspectives, drivers, and concerns to these GBN Strategic Planning themes:
    - → architectural transformation
    - → human health and comfort
    - → green construction management and productivity
    - → management, function, sustainment, and operation of buildings and users.
  - Use advanced data-driven tools for rapid GBN scenario generation and evaluation of multiple KPIs, such as Energy Consumption, Indoor Thermal Comfort, Air Quality, Investment, etc.
- Provide tools assisting the GBN initiators to produce a GBN Strategic Vision and KPI formalisation.
- Provide data-driven tools (DGNB) for rapid GBN scenario generation and evaluation.
- Scope the transition roadmap(s) from the Target Model(s) detailing the current maturity for each Living Lab, from which innovation testing and evaluation within the Digital Twin(s) can be executed, prior to physical testing and use.
- Define the GBN Integration Strategies, to manage and sustain GBN Transitions aligning to governance, people, process, technology, and information needs of a connected Europe, society, and environment. These tools and the decision support system will be refined in the LLs around the local policy to practice the context of individual development needs, interconnected with wider regional, national, and international strategies and trends.

WP2 aims to maximise the adoption of the PROBONO energy efficient approach and innovations beyond the project lifetime by:

- raising awareness of the PROBONO activities, results, and positive impact
- iteratively integrating stakeholder needs and requirements from the project onset
- validating the project findings against user needs and the market.

It will achieve this using a range of participatory and co-design methods that promote stakeholders (and citizen) participation.

Information gathered for WP1 and WP2 isn't technical, as it is related to the way the buildings are going to be used and how people are going to travel to the GBN.

The greater focus of WP1 and WP2 goal is to qualify and quantify. Therefore, it is highly difficult at this stage of the project to give a clear vision of what kind of data will be generated, used, or collected by them.

More information will be available in the DMP V2 (M30).

3.2 WP3

The aims of WP3 are to "provide GBN Models Innovations for Smart Green Building Construction and Renovation and SOTA Reports. Deliver a focused set of innovative solutions matching initial Living Labs requirements, which will be matured through application and evaluation in the Living Labs. These innovative solutions are GBN Insulation and green and cool roof-centric innovations, GBN related Construction and lifecycle blueprints, processes and controls including robots GBN Building Materials / upcycling and Social distancing (design for crowds, sensors) considering epidemiology risks. Identify systemic GBN innovations combining initially WP1 and WP2 findings and subsequently from knowledge gained from Living Labs (LLs) operation" (stated in the Grant Agreement).

	Data / Task 3.1	Data / Task 3.2	Data / Task 3.3	Data / Task 3.4	Data / Task 3.5	Data / Task 3.6
Data Origin	WP3 partners / technology providers	Task 3.2 partners / LLs	Task 3.3 partners / LLs	Task 3.4 partners / LLs	Task 3.5 partners/ LLs	Task 3.6 partners/ LLs
Data Type	Information on WP3 technologies and innovations	1) Information on Task 3.2 technologies and innovations 2) Information on the infrastructure, regulations and environment of the LLs*(see comments for more details)	1) Information on the approach being developed related to construction and life-cycle blueprints, processes and controls 2) Materials inventory for LCA in LLs	1) LLs regulations and climate conditions regarding the use of recycled materials, byproducts e.g asphalt 2) Technical properties on recycled materials and products	Regulations and measures regarding social distancing and epidemics	Data on the physical infrastructure, human behaviors, electrical demand at the GBN level resulting from buildings energy consumption / prices / renewable electric energy generation in LLs, weather data, existing charging infrastructure. Power generation mix etc

Data Purpose	Development of innovation models and DT models	Implementation of technologies in LLs / advance from TRL6-8	Record and analyse the lifecycle impact of options related to renovations and new constructions	Design, test and define new advanced products to satisfy requirements, regulations and high standards	Identify status, impact, SOTA and blueprints for tackling infection risks	Agent-based modelling of occupancy and mobility at GBN level
Which partner will use the data?	Task 3.1 partners	Task 3.2 partners / LL partners	Task 3.3 partners / LL partners	Task 3.4 partners	Task 3.5 partners/ LLs	Task 3.6 partners/ LLs
Data Processing	TBD	Design, Simulation and optimisation	LCA and BIM aided analysis and optimisation	Analysis and testing	Analysis and synthesis of data across representative buildings and infrastructure in each LL	Analysis and simulation
Data Format	Reports	Technical data sheets / Reports	Data sheets and reports	Data sheets and reports	Technical data sheets / Reports	Time-series data sheet, databases, technical reports
Public/ Confidential	Certain parts may be confidential	TBD / some parts may be confidential	TBD / some parts may be confidential	TBD / some parts may be confidential	TBD / some parts may be confidential	TBD / some parts may be confidential
Data Storage (Location, Duration)	MS teams dedicated project folder	MS teams dedicated project folder	MS teams dedicated project folder	MS teams dedicated project folder	MS teams dedicated project folder	MS teams dedicated project folder
Data harmonisation activities required	As per the deliverable D3.1 structure	TBD	TBD	TBD	TBD	TBD
Data merging activities required	TBD	TBD	TBD	TBD	TBD	TBD

3.3 WP4

The aim of WP4 is to provide a methodology and an internet-based planning tool to develop and calculate an optimized design for carbon neutral energy systems for GBNs, and to support all LLs in their implementation with the goal of achieving at least one carbon neutral green building, with a high share of locally generated renewable energy and a high efficiency per LL by the end of the project. Deliver a focused set of innovative solutions matching initial LL requirements, which will be matured through application and evaluation in the LLs, a GBN demand and response platform, Building Integrated Photovoltaics, and PROBONO GB Positive Energy Package.

For now, we do not have sufficient information regarding the data that will be collected, generated or created by the WP4.

3.4 WP5

The goal of WP5 is to produce the GBN Digitalisation Platform (GBN-DP), including the GBN Digital Twins (DTs), which will be interoperable representations of GBNs in the digital realm.

The GBN-DP will provide a structure to the components of GBN's in a format that is possible to interrogate, allowing decision makers to identify the most pertinent:

- Perform requirements elicitation for the different stakeholders and for the different GBNs lifecycle phases

- Design and implement the infrastructure supporting the GBN-DP

- Define and construct information and simulation models while ensuring their interoperability & integration

- Create a framework for tools and services development exploiting the GBN-DP

- Create a governance, access and security framework for data and services

	Meeting minutes	Recorded meetings	Source Code
Data Origin	WP5 Meetings	WP5 meetings	T5.2, T5.3, T5.4
Data Type	Documentation, reporting	Videos	Source Code
Data Purpose	Minutes of the WP5 meetings	• writing of Minutes of Meetings • replay of the meeting for partners who could not attend the meeting	Result of the software development
Which partner will use the data?	WP5 partners	WP5 partners	Tasks partners Living Labs where the tools will be deployed

Data Processing	Meeting recordings are stored in MS Teams allowing researchers to re-listen to understand a specific task.	Transcription of the video content to text for the meeting minutes	Compilation
			Packaging
Data Format	MS Word, PDF	MP4	Text
Public/ Confidential	Confidential	Confidential	Confidential
Data Storage (Location, Duration)	MS Teams for duration of project only	MS Teams for duration of project only	Source code repositories: Gitlab, Github repositories
Data harmonisation activities required	N/A	N/A	N/A
Data merging activities required	N/A	N/A	N/A

3.5 WP6

WP6 will use WP5 data (digital twins data) to calculate all the evaluation KPIs, but as this input data will be managed by WP5, WP6 will just consider the KPIs in the scope of the DMP.

WP6 will send to WP5 the formulas and algorithms to be used to calculate those KPIs, and all of these will be calculated within the digital platform to be developed in the scope of WP5.

	Main KPIs – PROBONO impacts	Auxiliary KPIs (support KPIs for Main KPIs calculation)	KPIs concerning exploitation, communication and dissemination activities	KPIs concerning WP2 social innovations	KPIs concerning WP3 technical innovations	KPIs -> WP4 tech. innovations	
Data Origin	Those KPIs will be calculated with data coming from WP5 (monitoring data)						
Data Type	KPI						
Data Purpose	Calculate all the KPIs needed to deploy the assessment activities and therefore to know the effectiveness of impacts achieved in each of the Living Labs once the innovations have been implemented						
Which partner will use the data?	Mainly WP6 and WP7 partners						

Data Processing	KPIs will be anonymised or protected. We need to know the building or the dwelling to which each subset or KPIs come from (the same with the monitoring information that will be used to calculate them (information coming from WP5))
Data Format	Mainly excel sheets and word
Public/ Confidential	Confidential in the beginning, but aggregated information not linked with specific buildings/dwellings could be published
Data Storage (Location, Duration)	Teams and in WP5 digital platform
Data harmonisation activities required	N/A
Data merging activities required	N/A

3.6 WP7

The Living Labs GBN Implementation will validate the PROBONO innovations under at least 2-3 operational environment and European climates and will demonstrate the via the Living Labs renovation scenarios the PROBONO capabilities via the technologies, the Digital Twin, and the monitoring framework. WP7 will provide both an experimentation and innovation environment and testbed for GBN solutions in six Living Labs.[2]

	Meetings	Communication Material with the LLs	Consent forms	WP7 Deliverables
Data Origin	Recording & minutes, files & information exchanged	Questionnaires, e-mails, workshops, meetings, documents	Beneficiaries & external participants	WP7 Contributors
Data Type	Documentation and reporting, recordings and internal monitoring	Documentation and reporting and internal monitoring	Paper signed document, Electronic signed document	Documentation and Reporting

---

[2] Source : PROBONO Grant Agreement - WP7 objectives

Data Purpose	WP follow-up and recording for those who cannot join the meeting	Providing Data from each LL to every interested party	Consent of partners' interviewed for tasks' purposes	
Which partner will use the data?	Partners to whom it may involve and the Management team	Partners to whom it may involve and the Management team	Partners whom it may involve (I.e., LLs leaders)	WP1-WP9 partners
Data Processing	Recording and minutes stored on TEAMS		yes	
Data Format	Word (.doc/.docx), Excel (.xls/.xlsx), PDF, MP4 videos	Word (.doc/.docx), Excel (.xls/.xlsx), PowerPoint (.ppt, .pptx) PDF	Word (.doc/.docx), Excel (.xls/.xlsx), PDF	Word (.doc/.docx), Excel (.xls/.xlsx), PowerPoint (.ppt, .pptx), PDF, dwg etc
Public/ Confidential	Confidential	Confidential	Confidential	Confidential
Data Storage (Location, Duration)	Microsoft TEAMS for the duration of the project + on folder/DB of WP 7 Partners	Microsoft TEAMS for the duration of the project + on folder/DB of WP7 Partners	Password secured folder/DB/platform or non-networked machine of the LLs leader or other partners undertaken the interview.	Microsoft TEAMS for the duration of the project + on folder/DB of project Partners
Data harmonisation activities required	N/A	N/A	N/A	N/A
Data merging activities required	N/A	N/A	N/A	N/A

3.7   WP8

The objective of WP8 is to raise awareness about the project achievements and facilitate the uptake of its findings as part of an integrated project strategy to maximise its impact. WP8 implements the dissemination and communication actions, using the PROBONO framework.

For now, we do not have sufficient information regarding the data that will be used, created or produced within WP8.

They will be more likely to be administrative data, but also dissemination and communication materials (use of PowerPoint and word for example, but also tweets, newsletters etc…).

3.8   WP9

WP9 aims to assess IPR and TRL status of exploitable results and continue the process of their characterisation. Create commercialization, sustainability, and replication strategies in support of the potential commercialisation of project outputs, maximizing the potential commercial value to partners in the EU and beyond, ensuring the financial viability on the longer run beyond the project. Set up and consolidate strategic contacts with exploitation partners and users. Accelerate the market uptake of PROBONO´s innovations by matching them directly to potential beneficiaries. Transfer project findings into standardisation activities and thus support the transfer of the project's findings to the public.

	Market information	Stakeholders' insights (personal vision on market information: needs, pains, expected gains, etc.)	Standards, legislation, and similar related information	General WP information (deliverables, project reports, partners contact details, meetings recording, partners information etc.) - similar to other WPs	Deliverables or publications
Data Origin	WP9 participants and publicly available data sources (i.e. market research reports)	Desktop research, stakeholders' interviews, Questionnaires, meetings and workshops	Public available sources	WP9 activities	WP9 activities
Data Type	Documentation, reporting, statistical data.	reports, questionnaires, etc.	Standards, laws, etc.	Documentation, reporting and internal monitoring	Documentation, reporting, statistical data.

Data Purpose	Develop, define and finetune a framework for assessing future exploitation options of PROBONO results.	Develop, define and finetune a framework for assessing future exploitation options of PROBONO results.	Define future replicability scenario of PROBONO results.	Monitoring and project planning	Reporting project progress and facilitate exploitation of project results
Which partner will use the data?	T9.1, T9.2 and T9.3 leaders	T9.1, T9.2 and T9.3 leaders	WP9 task leaders + MM	WP9 contributors and PROBONO management team	WP9 contributors
Data Processing	Based on project workplan	Based on project workplan	Based on project workplan	Based on project workplan	Based on project workplan
Data Format	MS Word, MS PowerPoint, MS Excel, pdf	MS Word, MS PowerPoint, MS Excel, pdf	pdf	Multiple formats	pdf
Public/ Confidential	Confidential	Confidential	Public	Confidential	Confidential and public
Data Storage (Location, Duration)	MS Teams for duration of the project	MS Teams for duration of the project	MS Teams for duration of the project	MS Teams for duration of the project	MS Teams for duration of the project
Data harmonisation activities required	N/A	N/A	Research for replicability framework	N/A	N/A
Data merging activities required	Expected generation of project deliverables	Expected generation of project deliverables	Expected generation of project deliverables	N/A	N/A

3.9   WP10

The aim of WP10 is to perform the overall management of the project to ensure smooth administrative and technical coordination, including procedures to ensure the completion of all deliverables in time and scope, within the budget and of the required quality. Administrative

and technical management, innovation management, quality, risks, and security management will be carried out to the full satisfaction of the EC and the Consortium Bodies.

This means that the WP10 will likely use Word (.odt/.docx), Excel (.xls/.xlsx), PDF, MP4 videos, PowerPoint, and will store information in MS Teams during the duration of the project.

4   Methodologies and Standards

This section covers the methodology and standards that will be applied to the project. The first section will present the FAIR principles, a guideline regarding data, and the second section will present the GDPR that concerns the data protection.

4.1   FAIR principles

The FAIR Principles refer to a concise, domain-independent, high-level and measurable set of guiding principles and practices that apply on a wide range of scientific data or metadata. They are the result of the work of a community of stakeholders representing academia, industry, funding agencies, and scholarly publishers in 2014, which was then adopted the same year by the European Commission as the data guidelines for the Horizon 2020 framework programme. They put "specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals."[3] The term "FAIR", refers to the characteristics of data or metadata of being Findable, Accessible, Interoperable and Reusable. In practice, the elements of the FAIR principles are related, but are independent and separable. Any combination of the principles can be applied incrementally. Thus, this modularity of the principles, as well as their distinction between data and metadata, facilitates their support in a wide range of special circumstances. The FAIR principles can also be applied to non-data assets which need to be identified, described, discovered, and reused in the same manner as data.

These principles constitute a general guide to the ''FAIRness'' of data. However, they themselves are not a standard or a specification. To be more precise, they precede implementation choices and do not necessarily suggest any specific implementation solutions. Instead, they act as a guide for data implementers, publishers, and managers to evaluate whether their particular implementation choices are rendering their digital research artefacts ''FAIR''. They form the basis for a long-term care of valuable digital assets composed by the data produced by the research project, while keeping the goal of being discovered and re-used by other research.

Given the definitions of ''FAIRness'', there is however a clear distinction between FAIR data and Open data, as the former does not necessarily imply the latter.[4] Indeed, while the openness of data is encouraged within Horizon 2020 programme, there are necessary and legitimate reasons to restrict access to certain data. Nonetheless, the ''FAIR'' principles can still apply equally to restricted data or internal data of an organisation, in order to make them more usable and of

---

[3] https://www.nature.com/articles/sdata201618

[4] https://ec.europa.eu/info/sites/default/files/turning_fair_into_reality_0.pdf

greater value. Following the principle of "as open as possible, as closed as necessary"[5], research data should be open by default, while setting a variable degree of openness. As illustrated by Figure 1, the more data becomes both open and ''FAIR'', the higher the benefits they bring.
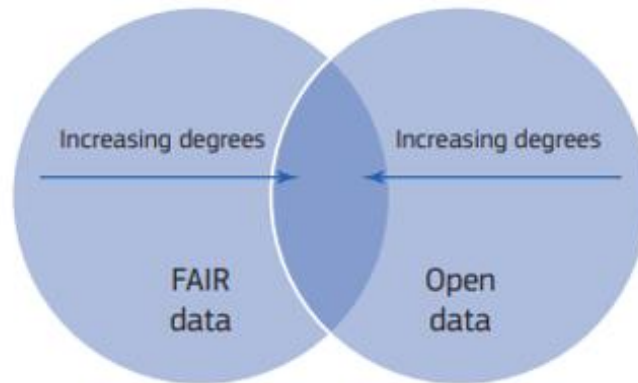

Figure 3: The relationship between FAIR and Open[6]

In general terms, the research data produced within Horizon 2020 should be ''FAIR''.

In the following sub-sections, the guiding principles of each one of the four concepts of ''FAIR'' are exposed. We are also detailing how PROBONO will fulfill the requirements of ''FAIR'' principles.

4.1.1   Making data Findable

To be ''Findable'' means:

F1. (meta)data is assigned a globally unique and persistent identifier
F2. data is described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly includes the identifier of the data it describes
F4. (meta)data is registered or indexed in a searchable resource

PROBONO as an H2020 project is obliged to provide a continuously updated Data Management Plan (DMP) that describes what data the project will use and produce, whether and how the data produced will be exploited or made (openly) accessible for verification and re-use, and how the data will be curated and preserved after the end of the project. Open data must be put into a public (research data) repository and for this matter PROBONO will use Zenodo (Figure 4). The repository should be OpenAIRE-compliant to enable the harvesting of metadata.

---

[5] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

[6] European Commission Expert Group on FAIR Data, Turning FAIR into reality, 26 November 2018, https://publications.europa.eu/en/publication-detail/-/publication/7769a148-f1f6-11e8-9982-01aa75ed71a1
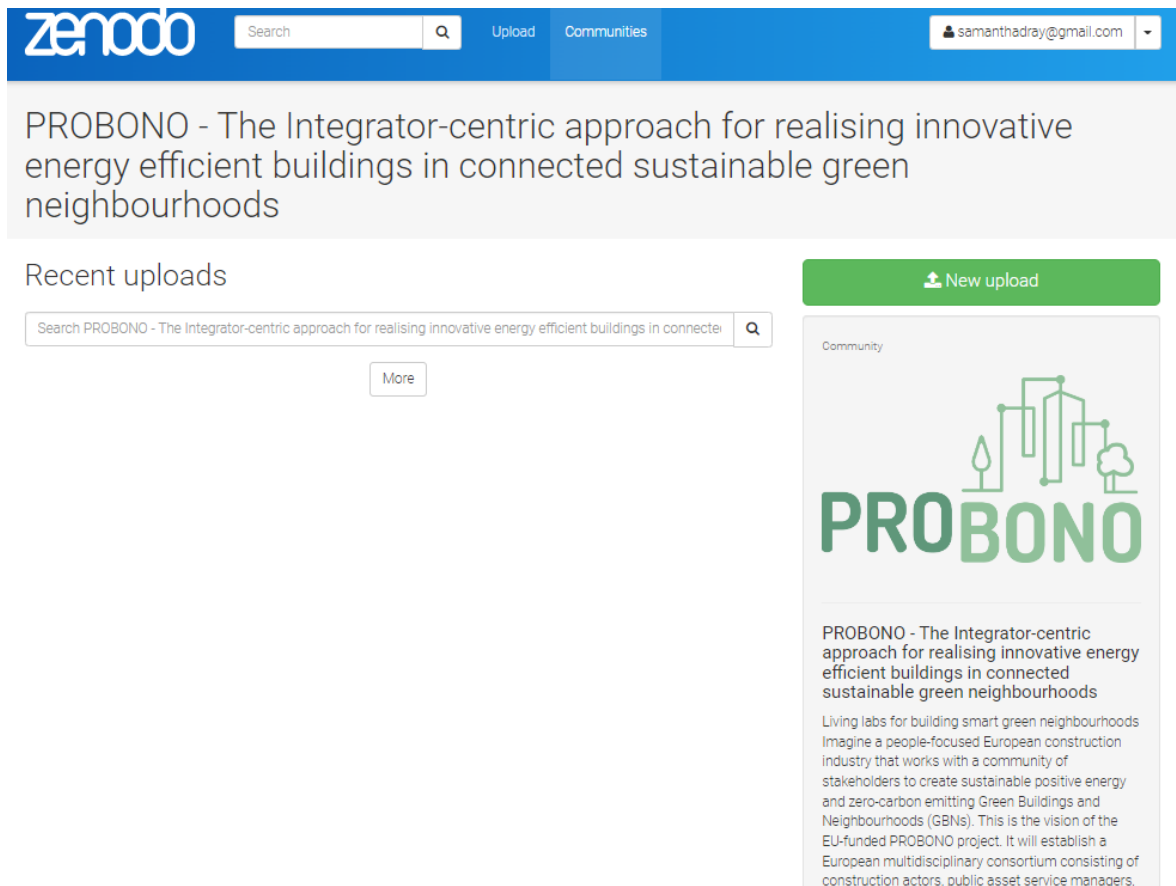
Figure 4 - PROBONO Zenodo repository

4.1.2   Making data Accessible

To be ''Accessible'' means:

A1. (meta)data is 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 is accessible, even when the data is no longer available

In compliance with the Horizon 2020 rules regarding Open Access (OA) (cf. 4.1 *infra*) to scientific literature, PROBONO will make any scientific publication accessible online for free under the scope of the project. We will choose the most suitable approach (either "green" OA or "gold" OA) to peer-reviewed scientific publications that might result from the project. The publisher will be chosen amongst those who respect both the authors' interests and accept the terms of open access publication (with an embargo period). On one hand, research data will benefit from an open access in a specific part of the project website, tailored to different levels of internal and external stakeholders. On the other hand, Partners will use an open access repository that will be connected to the tools proposed by the EC (e.g. OpenAIRE), in order to grant access to the publications and bibliographic metadata in a standard format, including the information requested by the EC.

4.1.3 Making data Interoperable

To be ''Interoperable'' means:

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

As data is still being produced, the strategy used to make them interoperable will be detailed in future versions of the DMP.

4.1.4 Making data Reusable

To be ''Reusable'' means:

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

To render data reusable, the PROBONO consortium will discuss how to add value to the data. The vision will be more precise in the DMP V2 (M30) as we do not have sufficient information at this stage of the project.

4.2 GDPR

Within the European Union, during the early stages of the internet, the 1995 Data Protection Directive was adopted. Over the last 25 years, tremendous changes in technology brought forth the need to revise data protection regulations. In 2016, the EU adopted the General Data Protection Regulation (GDPR), which is now recognised as law across the EU.

GDPR has been enforced in 2018 and provides broader data privacy for individuals, as well as new obligations for any data-processing organisation (data collector, publisher, implementor), regardless of geographical location, that collects, uses, or processes European Union citizens' personal information. Personal data security is thus improved via the enforcement of the four following aspects.

- **Right to be forgotten**: Article 17 mentions the obligation of the organisation, and any business partners with whom they have shared data, to delete any personal data from their systems upon request.

- **Data protection by design and default**: Article 25 stipulates that organisations must set internal policies and measures to protect data by design and default, and that all applications, services, and products must adhere to these policies.

- **Secure data processing**: Article 32 requires that organisations be able to prove an implementation of measures to ensure appropriate levels of security.

- **Timely breach notifications**: Article 34 imposes hefty fines to organisations if breaches of unencrypted data are not reported to authorities and affected individuals within 72 hours.

GDPR has been identified as one of the targets of PROBONO. In reality, the current context of the Project shows a variety of factors bringing different barriers and obstacles for it to reach higher TRL. Amongst the legal factors, GDPR plays a data protection role and its purpose limitation (requirement that personal data be collected for specified, explicit, and legitimate purposes, and not be processed further in a manner incompatible with those purposes) can hinder our progress. For this reason, PROBONO will anticipate these limitations thanks to legal studies in WP10, and include an assessment of factors influencing data governance, including applicable legal regimes such as the GDPR.

4.2.1   Personal Data Collection, Processing, and Re-use

A limited number of PROBONO partners will be involved in personal data collection and processing. They will also be involved in tracking or observation of participants arising from Datasets and in further processing of previously collected personal data (i.e. secondary use). PROBONO will be subject to the provisions of the GDPR that enshrines the following key principles (without considering exemptions):

1. Data must be processed fairly, lawfully and only for the purpose for which it was collected and further processed

2. Data cannot be disclosed without authorisation unless there is an overriding act of law or legitimate grounds to do so

3. Subject to certain exemptions, individuals have a right to access the information relating to them and to ask for correction of inaccurate data

4. Information cannot be transferred beyond the EEA boundaries without consent or adoption of other adequate protection measures

5. Organisations are usually required to register or notify the processing of personal data unless the data processing is simplistic, or a data protection officer has been appointed

6. Organisation must have adequate security measures in place

7. The PROBONO consortium exercises a principal of minimum resort to notification exemptions afforded under national laws

4.2.2   Roles and Responsibilities

Data management activities concern the whole project and need to be coordinated and monitored both at project and work package level. Data management is also linked to publication of project results and thus dissemination activities. Therefore, the following roles and responsibilities can be identified:

➔ The Data Protection Officer (T10.4 task leader) is responsible for:

- Developing the data management plan and policy in cooperation with the project management in WP10 and the technical partners

- Developing a user guide for the usage of PROBONO's living DMP

- Advise data controllers and processors within the PROBONO project on the processing of personal data, training of researchers and assignment of responsibilities

- Assist in risk assessment of personal data processing

- Monitoring data management activities (both collection and publication) and deadlines and sending reminders to WP data managers

- Providing support to WP data managers

- Coordinating the writing of the DMP deliverable documents (D10.11-12-13)

- Providing solutions for specific issues in accordance with project management

- Cooperate with any national or European supervisory authority and act as contact point for the project with such authorities

➔ The Work Package Data Managers are responsible for:

- The implementation of the data management policy in their respective WPs

- Monitoring data management activities and deadlines and sending reminders to partners

- Offering customized help and further guidance for using PROBONO's living DMP

- Asking partners for missing information or clarifications

- Providing input to the DMP deliverable documents (D10.11-12-13) by analysing and summarising the WP- specific datasets listed in PROBONO's living DMP

- Offering customized help and further guidance for publishing open data and open-source software

- Monitoring that open results (data and software) are deposited in the default repository or a complementary OpenAIRE-compliant repository and sending reminders to partners

- Monitoring that open results available in OpenAIRE are properly linked (OpenAIRE PROBONO) with PROBONO

- Contacting the quality assurance and ethics committee in case of questions and ethical and privacy issues that may forbid a publication of the data

- Ensuring that the meta-data of data used and produced at Work Package-level is made available in PROBONO's living DMP according to the PROBONO data management policy and guidelines in a timely manner.

Samantha Dray acts as the point of contact for Data Protection Issues and Data Protection Officer (DPO) in the project.

WP Data Managers will be named inside each Work Package.

4.2.3   Compliance with GDPR Recital 78

Personal Data within the scope of the Directive. It is necessary for the data to be related to an identified or identifiable living individual. The individual need not be directly identifiable but may be identified by a reference number or some other tag which, in a given small group or through analysis of patterns in sufficient volumes of data, might allow an individual to be singled out from a group. Based on the kinds of data sources to be included in this research, direct personal identifiers (e.g. specific names or faces) may exist in a variety of locations within the dataset. PROBONO's default anonymisation process(es) will be 'one-way,' with original source data being disposed, so that re-identification of data or decoding of anonymisation tokens by reference to

any 'real-world' data sets will be rendered difficult to the greatest extent possible. PROBONO will follow the guidance set forth in the Article 29 Working Group 05/2014 Opinion on Anonymisation Techniques[7], specifically its recommendations on Pseudo-anonymisation, Noise addition, Substitution, Aggregation, K-anonymity, L-diversity, Differential privacy, and Hashing/Tokenization. PROBONO will also include downstream contractual obligations as a legal measure to respecting privacy in the use of the project results.

4.2.4   Data Protection Agency Notification

The data processed in PROBONO is unlikely to constitute Personal Data within the meaning of the EU Directive and relevant national legislation. PROBONO is also likely to be exempted from national notification processes because our data collection is for the purposes of scientific research and, thus additional institutional data protection measures, access restrictions, etc. will be put in place. We are, nevertheless, mindful that anonymisation approaches must be applied to video/still images within data to avoid the risk that a token identifier might become associated with sufficient unique data points to uniquely identify a living individual. We also make sure to notify data protection authorities in jurisdictions where research activities will be carried out and specific relevant actions within it in order to obtain (if necessary) authorisation for such activities. The exact requirements and due diligence will need to be scoped and defined within the relevant jurisdictions.

The relevant national approvals will be sought and acquired when necessary, as the notification requirements vary from one country to another, and therefore no single timeline can be provided for completion of all notification procedures.[8] Renewal of notifications, when necessary, will be carried out in line with requirements of different national legislations. Processes for notification vary from one jurisdiction to another. The following project principals have been assigned the responsibility of acting as interlocutors with their own national data protection agency. Further information on notification procedures and the relevant agencies in Europe can be found in the *Article 29 Working Group document*s and the relevant agencies in Europe can be found in the *Article 29 Working Group document "Vademecum on Notification Requirements"*

The exact partner and contact person who will notify the relevant Data Protection Agency will be determined during the project and after the transition from *Directive 95/46 to General Data Protection Regulation (GDPR) 2016/679*.

Renewal of notifications, when necessary, will be carried out inline with the requirements of different national legislations. The renewal requirements will also be leveraged to react to change in the project and the adoption of video archives into research, development, and test tasks.

4.2.5   Compliance with Article 49 of the GDPR

While our position as scientific researchers permit us derogation from the prohibition on processing (sensitive categories of) Personal Data, we are nevertheless aware that it remains

---

[7] http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

[8] http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/others/2006-07-03-vademecum.doc.

incumbent upon us, to provide specific and suitable safeguards so as to protect the fundamental rights and privacy of Data Subjects. Some of these safeguards are already detailed above. PROBONO further ensures that any Personal Data collected will also be treated in accordance with Article 49. In particular, collected Personal Data will be processed fairly and lawfully and will be used only for research purposes as specified in our original proposal. The data will be adequate, relevant, and not excessive in relation to the purposes for which they are collected. We will endeavor not to collect and to expunge all data that is not directly project-related.

4.2.6   Safe Harbour and Privacy Shield Considerations

In light of the *Court of Justice of the European Union October 6th, 2015 decision on the EU-US Safe Harbour agreement*, the PROBONO project will store all data derived from personal data (after anonymisation or dissociation) in EU member states and comply with the *Article 29 Working Groups* communiques on transfer of data outside of the union and forthcoming member state decisions on Safe Harbour.

4.2.7   Activities Dedicated to Ethical Considerations

The guiding principles at the heart of the PROBONO approach are the highest ethical standards, the protection of privacy and the validity of data and its accurate representation. In adhering to these principles and remaining cognisant of concerns that arise in the Work Plan, PROBONO will take the following steps, in addition to those detailed above, toward addressing these:

- Compliance with the policy recommendations made in the Social Impact Expert Working Group EC DG ENTR Report (2012);

- The availability of partners with Ethics, Privacy and Legal expertise to all PROBONO project staff members at the outset of the project and throughout;

- Assurance that Privacy by Design, Ethics, Legal and Societal Impact requirements are included as research and development mandates integrated into the PROBONO project plan in compliance with GDPR Article 25.

- Test and evaluation of research results will be carried out in WP8 under the principal of informed consent.

4.2.8   Minimum Resort to Exceptions and Derogations

The GPDR allows for exceptions and derogations for personal data used for research. For example, general exemptions for processing of certain categories of sensitive personal data (e.g. Article 6 and Recital 50). Exceptions for a right to opposition for processing or storage of data (Article 89), and for processing of data without consent (Article 6.1.f, Recitals 47 and 157) may be applicable. PROBONO commits to a minimum resort to exceptions and derogations in the processing of personal data within the project for the purposes of research.

4.2.9   Activities Dedicated to Protection and Securing of Personal Data

As project coordinator, ACCIONA shall ensure the consortium guarantees adequate treatment of personal data generated during the project. This will be done via a set of development directives and methodologies. To ensure that security systems development principles are integrated from the inception of the project best practices will be issued to RTD staff in the PROBONO project to ensure the project applies adequate database encryption and secure systems techniques.

Furthermore, the directives and mandates will also integrate the technical requirements of European, national and regional data protection legislation. Partners will be required to have adequate security measures in place - both technical (firewalls, access controls, Access audits, etc.) and operational (training, incidence reporting, etc.) As both part of the project tasks, and notification procedures, the following range of issues will be considered in establishing such directives and mandates:

- Categories of sensitive data

- Security measures for sensitive data

- Policies for fair acquisition and processing of data

- Data retention policies

- Legal Basis for the information processing

- Policies for processing compatible with purpose

- Polices for Data Controller and Data Processors

- Description of the technical characteristics of the data processed

- Technical features and topology of the information systems where data is stored and processed

Further consideration shall be given to the following relevant regulation, decisions and guidelines. The following EU regulations are recognised to be relevant for the project:

- *The Charter of Fundamental Rights of the EU (especially articles 3, 7, 8 and 25).*

- *Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use.*

- *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.*

- *Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention Nr. 108).*

- *WMA Declaration of Helsinki (especially articles 13, 20, 21, 22, 23 and 24, 25).*

- *Convention of the Council of Europe on Human Rights and Biomedicine signed in Oviedo on 4 April 1997.*

- *CIOMS -- International Ethical Guidelines for Biomedical Research Involving Human Subjects.*

- *Code of Nuremberg (Article 10).*

- *Ethics and EU funded research Council Decision 1513/2002/EC on FP6 (e.g. article 3).*

4.2.10   Shared Information and Personal Data

The Parties agree that any Background, Results, Confidential Information and/or any and all data and/or information that is provided, disclosed, or otherwise made available between the Parties during the *implementation of the Action and/or for any Exploitation activities ("Shared Information"), shall not include personal data as defined by Article 2, Section (a) of the Data Protection Directive (95/46/EEC) (hereinafter referred to as "Personal Data") or under Article 4.1 of the GDPR. Accordingly, each Party agrees that it will take all necessary steps to ensure that all Personal Data is removed from the Shared Information, made illegible, or otherwise made inaccessible (i.e. de-identify) to the other Parties prior to providing the Shared Information to such other Parties*.

5   Allocation of resources

In the project, Samantha Dray Ait Melil (AKKA) plays the role of Data Manager and liaises with the Technical Management Team (TMT) about the data management issues. The Data Manager leads data management plan tasks and participates in the project coordination in terms of the evaluation data collection, storage and handling, as well as their publication as part of the ORDP.

All research data collected as part of this project is owned by the data producer or Partners involved in trial sites. The Partners in PROBONO will take the responsibility for the collection, management, and sharing of the research data. Quality assessment will be the responsibility of data manager of each trial site.

The costs to make the data ''FAIR'' in PROBONO shall be handled by each partner who will have to generate its data according to the requirements expressed in the Data Management Plan. Specifically, different tasks of WP10 will contribute to fulfil the requirements of FAIR principles.

6   Access and Curation

This section explains, in the first part, the strategy chosen about the access of the data (storage, access policies, licensing). The next part covers the data protection and qualification.

6.1   Data Access

6.1.1   Open Access and ORDP

Within Horizon 2020 (H2020), the objective is to have an open access to scientific information. The notion of open access here refers to "the practice of providing online access to scientific information that is free of charge to the end-user and reusable" [9], where "scientific information" includes both the peer-reviewed scientific research articles/publications and the research data underlying publications, curated data, and/or raw data.

---

[9]   https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

To ensure open access to all peer-reviewed scientific research publications relating to project results, which is an obligation, two routes are possible.

- Self-archiving / 'green' open access - the author, or a representative, archives (deposits) machine-readable electronic copy of the published version or the final peer-reviewed manuscript in an online repository before, upon or after its publication. Some repository software delays access only after an embargo period has elapsed. If this is the case, the European Commission demands that the open access is ensured within a maximum of six months.

- Open access publishing / 'gold' open access - an article is immediately published at a publisher or on a journal website. In this model, open access must be granted at the latest on the date of publication. A copy of the publication should also be deposited in a repository.

Whichever the route chosen, H2020 beneficiaries must at least ensure that any scientific peer-reviewed publications can be read online, downloaded and printed. They must also strive to provide the right to copy, distribute, search, link, crawl and mine to the public, in order to make publications more useful. As mentioned previously (cf. 3.1.2 supra), peer-reviewed scientific publications resulting from the project will become accessible openly, thanks to an open access repository used by partners, connected to the tools proposed by the EC (e.g. openAIRE), which grants access to the publications and bibliographic metadata in a standard format, including the information requested by the EC.

With regard to the openness of research data, the Open Research Data Pilot (ORDP) run by the European Commission "aims to improve and maximise access to and re-use of research data generated by H2020 projects and takes into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions." This ORDP applies primarily to the digital form data needed to validate the results presented in scientific publications, including statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. Users can normally access, mine, exploit, reproduce and disseminate openly accessible research data free of charge.

Despite the benefits of open access, some research data cannot be made open. For this reason, the principle of "as open as possible, as closed as necessary" applies. It is therefore "possible to opt out of research data sharing at any stage - before or after the signature of the grant agreement - but reasons have to be given e.g. for intellectual property rights (IPR) concerns, privacy/data protection concerns, national security concern, if it would run against the main objective of the project or for other legitimate reasons.", precisely the same potential reasons which may play a role in the balance.

In order to specify the type of data subject to open access, a Data Management Plan (DMP) is required, which is the subject of this present deliverable. The management of scientific information within H2020 under the open access guidelines is illustrated as follows.
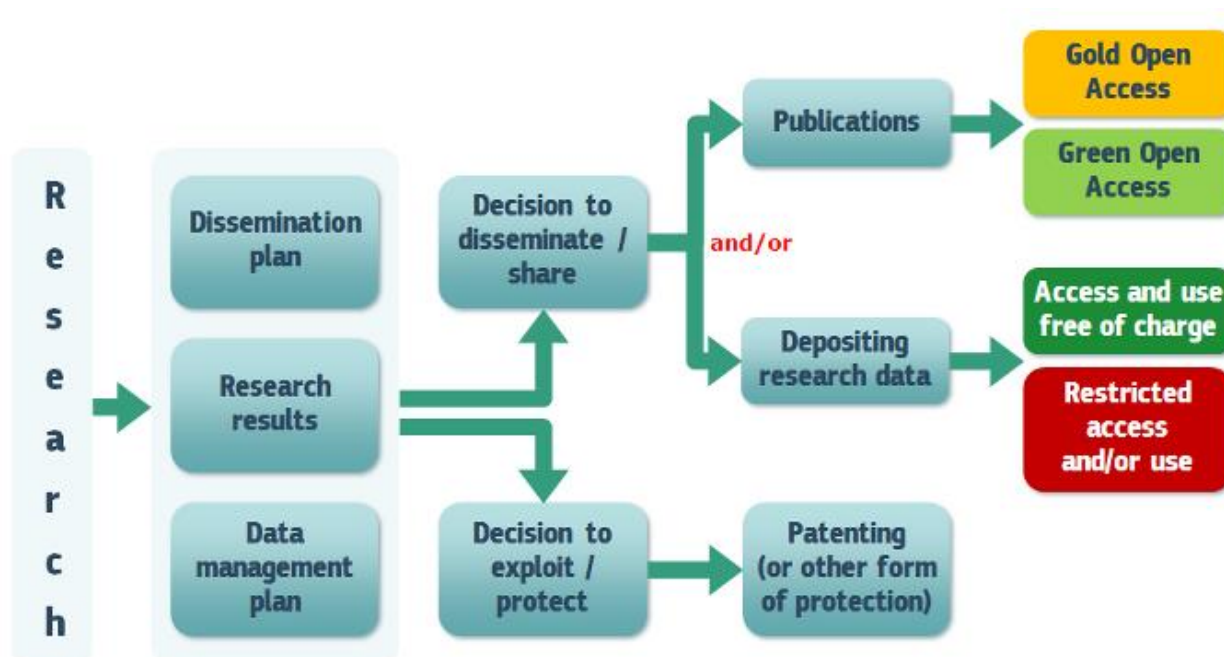
Figure 5 Open access to research data[10]

As mentioned in the Grant Agreement, PROBONO is committed to ORDP.

As already explained previously, the two pillars to ORDP consist of "the development of a Data Management Plan (DMP) and the open access to research data, if possible"[11]. The first point is being addressed by the establishment of a DMP as a deliverable of PROBONO and by its maintenance. The second point will be fulfiled by the open access to research data including public deliverables, non-protected results or results coming after a patent registration, which will be rendered throughout the whole project's duration in a specific part of the project website, tailored to different levels of internal and external stakeholders but also published on a general purpose open repository like Zenodo. The partners will agree altogether on the data that are shareable and the ones that are confidential and that should remain in the consortium only.

6.1.2   Data Licensing

The results of the PROBONO project, i.e. data, knowledge or information produced and generated is attached to the intellectual property rights pertaining to their usage. In absence of specification, national intellectual property rights apply. However, national intellectual property rights are different from one country to another, thus it is difficult to interpret voluminous national laws of several countries. Short terms of use adapted for the special case are usually the better solution.

---

[10] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

[11] https://www.openaire.eu/what-is-the-open-research-data-pilot

In order to increase data reuse on the results of the project, Partners should assign a license to their shared data, so that results can be exploited. A license describes the conditions under which the data or software is (re)usable. The license should contain:

- a disclaimer of warranty and limitation of liability
- a citation requirement
- rules for databases and collections.

For example, Creative Commons Corporation (CC) has introduced a well-known system of licenses.[12] CC licenses are free of charge and easy to use since any license consists of a legally binding arrangement, a summary, a symbol which may be downloaded as a graphics file and inserted in a document, and a machine-readable symbol. Their public copyright licenses incorporate a unique and innovative "three-layer" design: Legal Code, Commons Deed and CC REL. This type of license may be utilised for the protection of all types of content. It also makes the concept understandable by the creators of works, uses and the Web itself.

Besides, Open Source Initiative (OSI) approves some licenses which comply with the Open Source Definition and which have gone through the OSI license review process. Such approved licenses allow software to be freely used, modified, and shared.[13] They include Apache, GNU, MIT, Mozilla, just to mention the most popular ones. The Consortium will choose the most appropriate licenses for use depending on the type of data concerned.

6.1.3   Storage

In the European framework, European Open Science Cloud (EOSC) is an initiative providing a world-class data infrastructure FOR storING and managING data. It arises from the EC's aim to promote the access and reuse of research data that comes out of publicly funded research. Under Horizon 2020, EOSC becomes the best instrument to provide a framework for collaboration and the pooling of resources at European, national, regional and institutional levels.[14] In particular, EOSC aims at solving the problem of fragmented access and non-interoperable research data centres across Europe. In terms of data storage, EOSC provides several reliable, secure and scalable cloud storage solutions for scientific data, apps and workloads. Amongst them, some are of open access or even fully open access. For example, BlueBRIDGE suggests an Online environment to support secure and controlled data sharing and storage, INRIA's Software Heritage archive is a great library of source code on engineering and technology sciences, and Institute of Informatics of Slovak Academy of Sciences offers EGI FedCloud client.

EOSC also provides other services such as Open Science (OS) publishing infrastructure which is connected to the EOSC. OpenAIRE is one such an example which contributes actively to EOSC.[15] On one hand, OpenAIRE provides a cost-free data repository (Zenodo), allowing data deposit and discovery, rendering data Accessible. On the other hand, through harvesting repositories

---

[12] https://creativecommons.org/licenses/

[13] https://opensource.org/licenses

[14] https://eosc-portal.eu/about/eosc

[15] https://www.openaire.eu/openaire-and-eosc

and mining techniques, OpenAIRE infers links between publications, research funding, and research data, which enables data Reuse.

As mentioned in the previous section concerning Open access, PROBONO partners will use such an open access repository as OpenAIRE, in order to grant access to the publications and bibliographic metadata in a standard format. Specifically, Zenodo is a repository where every upload is assigned a Digital Object Identifier (DOI), making them citable, trackable, and thus Findable. It also provides flexible Access conditions. Zenodo also ensures Interoperability by allowing itHub integration to preserve metadata. Finally, its versioning gives the possibility to update datasets easily and store all changes in metadata over record's lifetime, which enable Reuse. To summarise, these advantages are totally compliant with the FAIR principles.

6.1.4   Authentication and authorisation

During their storage, data should be protected against any type of modification by the implementation of some security principles. The security principles are listed below:

- Authentication: All the users wanting to get access to the PROBONO data servers should be authenticated. Also, proper means are used to authenticate the servers. An authentication system could be used to handle the authentication of the users during the project.

- Authorisation: the access to PROBONO data servers is only available to the authenticated and authorised users. These categories and the rights of those users are defined and enforced. The appropriate access control policies and mechanisms (including physical access control) shall be identified project wide and for each trial site. Those actions are necessary to ensure the security of the data.

- Accounting: In PROBONO, any access and modification to a resource by any user is securely logged in order to prevent users from denying that data files were accessed, altered or deleted. Other accounting mechanisms shall be implemented.

- Confidentiality: the data stored in PROBONO servers should be encrypted during transmission and storage.

- Communication Security: Access to PROBONO servers should be done through encrypted communication channels such as HTTPS and IPsec.

- Data Integrity: The data collected during PROBONO should be protected from malicious and accidental modifications by any users during their transmission or their storage. Cryptographic mechanisms such as hash functions and digital signatures shall be used.

- Availability: This security principle assures that the PROBONO servers should be available for PROBONO users during the defined interval of service. Also, regular backups of the data should be made. Therefore, mechanisms to cope with the charge and DoS attacks should be implemented.

6.2   Data qualification and curation

6.2.1   Data protection

Securing stored digital data involves preventing unauthorised people from accessing it, as well as preventing accidental or intentional destruction, infection, or corruption of information. While data encryption is a popular mechanism, it is just one of many techniques and

technologies that can be used to implement a tiered data-security strategy. Steps to secure data involve understanding applicable threats, aligning appropriate layers of defence, and continual monitoring of activity logs taking action as needed. This means that a multi-tier approach needs to be adopted from all the partners.

The proper method of storage and the appropriate community along with levels of access for privileged users are important considerations for comprehensive protection. Improperly stored information along with overly permissive accounts are a centralised theme in many high-profile breaches. Partners within PROBONO will follow a specific set of guidelines to comply with the project's main requirement for storage of digital data.

- Data availability must be guaranteed.

- Confidential data must be stored using access protection.

- Strictly confidential information must only be stored in an encrypted mode.

- Confidential data must not be stored in online services that are not approved by the PROBONO Consortium.

- Any exception from this measure must explicitly be approved.

- Modifications to data with high integrity requirements must be documented and approved by the partners.

6.2.2   Data qualification

The aim of the data qualification is to clarify and thoroughly document the semantics of the data, which facilitates subsequent data analysis. If there is an original data documentation, it typically focuses on what is of interest to the management or the auditors. Yet, data qualification allows to the exploration of the semantics of the data in detail, which then be updated or specified with this documentation. Indeed, data qualification allows to understand and reveal the gap between the perceived semantics of current data analysers and the true and underlying semantic of the data.

Data is dynamic and requires constant maintenance. Data qualification should thus be constantly carried out, so as to keep the data up-to-date, valid, accurate and reliable. In order to operate data qualification, additional computation is needed to implement rules for different processes of data verification, validation, approvals, and rejection, which then generates metadata. Although qualifying data necessitates additional resources and efforts, better qualified data delivers higher quality and consistency, which helps with decision making and clarifies the direction of further research.

Conclusion

This Data Management Plan V1 gives a first overview of the data processed in PROBONO and describes the data categories of the project, providing information on their management and on the implementation of the FAIR principles in the project. PROBONO will be committed to following the principles of the Horizon 2020 Open Research Data Pilot, and therefore the guidelines associated with open access have been described, with the aim to ensure that the project's outputs will be openly available to the research community.

Towards the Data Management Plan V2:

It is important to note that the Data Management Plan is a living document and will be constantly updated until the end of the PROBONO project. The present document is the first version of the Data Management Plan, and in the next version (D10.12, 2nd issue), more details will be provided, especially regarding the description of shared datasets, and corresponding standards and methodologies. A great consideration will be given to datasets to be shared in the course of the Living and Virtual Labs (WP7) and for which we do not yet have sufficient information at this stage of the project. These different categories of data require different treatments and the definition by the project of an appropriate methodology (for example anonymisation of personal traveller data). In the same way, the data from our reference group of stakeholders will be detailed, and their specificities duly described.

PROBONO consortium

The consortium of PROBONO consists of 47 partners with multidisciplinary and complementary competencies. This includes leading universities, networks and industry sector specialists.

@PROBONO
#PROBONO (https://twitter.com/probonoh2020)

https://www.linkedin.com/company/PROBONO

For further information please visit https://www.probonoh2020.eu/

PROBONO

The content of this document reflects solely the views of its authors. The European Commission is not liable for any use that may be made of the information contained therein. The PROBONO consortium members shall have no liability for damages of any kind that may result from the use of these materials.

References

[1] https://www.worldgbc.org/what-green-building

[2] European_Green_Deal_EN_200710_fin

[3] SET-Plan Action 3.2:
https://setis.ec.europa.eu/system/files/setplan_smartcities_implementationplan.pdf

[4] Internal Electricity Market Directive (EU) 2019/944 5 Renewable Energy Directive (EU)

[5] Renewable Energy Directive (EU) 2018/20012018/2001

[6] https://dk-gbc.dk/publikation/guide-to-dgnb-for-buildings

[7] https://op.europa.eu/en/publication-detail/-/publication/16cd2d1d-2216-11e8-ac73-01aa75ed71a1/language-en

[8] https://www.lifecycleinitiative.org/starting-life-cycle-thinking/life-cycle-approaches/social-lca/

[9] https://ec.europa.eu/environment/gpp/lcc.htm