



D6.2 - Data management plan

WP6, Task 6.2

June 2024 [M6]

Author(s): Inès Elisabeth KAHIN (R2M Solution), Sylvain Kubicki (LIST)

Disclaimer



regenproject.eu



@REGEN EU Project



@REGENEuProject



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

The content of this deliverable reflects only the author(s)' view. The European Union or CINEA is not responsible for any use that may be made of the information it contains.

Copyright notice

©2024 REGEN Consortium Partners. All rights reserved. REGEN is a HORIZON EUROPE Project supported by the European Commission under contract No. 101123325. For more information on the project, its partners and contributors, please see the REGEN website (www.regenproject.eu). You are permitted to copy and distribute verbatim copies of this document, containing this copyright notice, but modifying this document is not allowed. All contents are reserved by default and may not be disclosed to third parties without the written consent of the REGEN partners, except as mandated by the European Commission contract, for reviewing and dissemination purposes. All trademarks and other rights on third party products mentioned in this document are acknowledged and owned by the



Co-funded by
the European Union

respective holders. The information contained in this document represents the views of REGEN members as of the date they are published. The REGEN consortium does not guarantee that any information contained herein is e-free, or up-to-date, nor makes warranties, express, implied, or statutory, by publishing this document.

Document Information

Grant agreement	101123325
Project title	REGENeration of neighbourhoods towards a low-carbon, inclusive and affordable built environment
Project acronym	REGEN
Project coordinator	Sylvain Kubicki (sylvain.kubicki@list.lu)- Luxembourg Institute of Science and Technology
Project duration	1 st January 2024 – 31 st December 2027 (48 Months)
Related work package	WP 6 - Project Management
Related task(s)	Task 6.2 - Data management Plan
Lead organisation	R2M Solution
Contributing partner (s)	All partners
Due date	30/06/2024
Submission date	28/06/2024
Dissemination level	PU - Public

History

Date	Version	Submitted by	Reviewed by	Comments
27/05/2024	V0.1	Inès Elisabeth KAHIN (R2M)	Sylvain Kubicki (LIST)	First draft of the DMP
28/06/2024	V1	Inès Elisabeth KAHIN (R2M)	Sylvain Kubicki (LIST)	Final version of the DMP

Table of content

1. Executive Summary.....	6
2. Keywords.....	6
3. INTRODUCTION.....	7
1. Purpose and target group	7
2. Relation to other activities in the project	7
3. Report structure	8
4. APPROACH TO DATA MANAGEMENT	9
1. Open access	9
2. Open data.....	9
3. Data storage and sharing	10
4. Copyright licences.....	10
5. DESCRIPTION OF PROJECT DATASETS.....	11
1. List of collected datasets.....	11
2. The form - template	12
3. Data collection in WP1.....	14
4. Data collection in WP2.....	20
5. Data collection in WP3.....	39
6. Data collection in WP4.....	50
6. Conclusions.....	58

Abbreviations and acronyms

Acronym	Description
AFUR	Assessment Framework for Urban Regeneration
BIM	Building Information Modeling
CC	Creative Commons
DMP	Data Management Plan
DNSH	Do No Significant Harm
GA	Grant Agreement
GDPR	General Data Protection Regulation
OGC	Open Geospatial Consortium
RDI	Research Data Infrastructure
WP	Work Package

1. Executive Summary

This Data Management Plan (DMP) outlines how research data generated or collected during the course of the project will be managed, stored, and shared. It serves as a roadmap for handling data throughout the project lifecycle, ensuring compliance with Horizon Europe's requirements and promoting good data management practices in the REGEN project. It clarifies which data will be openly accessible and which will remain confidential within the consortium.

Following Horizon Europe guidelines, the report emphasises the importance of Open Access (OA), defined as the free online provision of reusable scientific information. The rationale behind OA is underscored, highlighting the benefits of openly sharing publicly funded research data with the research community, commercial sector, and civil society.

While certain collected data will be openly disseminated, certain work packages may generate internal-use data primarily intended for consortium management and communication.

Although most project datasets will be openly accessible, exceptions may arise due to privacy concerns, administrative burden, or limited utility to external users. Nevertheless, the project endeavours to openly disseminate its research findings, safeguarding against exploitable results or privacy infringements.

Specific disclaimer:

- The structure and the approach of this deliverable is common for data management deliverables written by R2M Solution in European projects and thus contains similar content as other Data Management Plan (DMP) deliverables.

2. Keywords

Data management plan, datasets, open access, open science, FAIR Principles (Findable, Accessible, Interoperable, Reusable)

3. INTRODUCTION

1. Purpose and target group

REGEN aims to decarbonize European neighbourhoods and cities by addressing socioeconomic, climate, and built environment challenges through a multifaceted Toolbox. This Toolbox will use cutting-edge digital technologies and life cycle sustainability assessments, engage citizens, and influence behaviour change. The project will validate a catalogue of 50 urban regeneration interventions with an Assessment Framework for Urban Regeneration. This framework will estimate the revitalization potential of urban infrastructure, using digital twins for monitoring and sustainability methodologies for evaluation.

The Data Management Plan (DMP) for the REGEN project outlines the systematic organisation, storage, and dissemination of data generated or collected throughout the project's duration. It delineates the procedures for handling both open access and confidential data within the consortium, adhering to relevant regulations and guidelines. The DMP ensures transparency and accountability in data management practices, facilitating collaboration and knowledge sharing among project partners. It also addresses data security and privacy considerations, safeguarding sensitive information while promoting open access to research outcomes wherever feasible.

This report serves two primary audiences. Firstly, it is targeted internally towards the consortium partners organisations involved in the project's implementation across Luxembourg, France, Ireland, Spain, Italy, Netherlands. The Data Management Plan (DMP) aims to establish consistent data handling practices among these partners to enhance efficiency and reliability throughout the project's execution.

Secondly, the report is directed towards a broader audience interested in the multifaceted Toolbox designed by REGEN. The DMP will delineate standard data formats, meaningful metadata, and open repositories to facilitate data sharing and enable widespread replication of the project's outcomes.

2. Relation to other activities in the project

As part of Work Package 6, the Data Management Plan (DMP) has been developed under the coordination of R2M, with significant input from other project partners. Work Package 6 is directly involved in executing the DMP, with each partner responsible for their respective activities. The most crucial datasets for validating the project's impact and drawing conclusions in scientific publications are the ones regenerated in WP3 and its 4 large-scale demos in Luxembourg, Spain, Ireland and

Italy, as well as T4.1 that will evaluate the demo sites' impact. These datasets will play a pivotal role in affirming the project's outcomes and informing future research endeavours.

3. Report structure

The report begins by outlining its purpose, intended audiences and the process for ongoing development. Section 2 outlines the concepts of open access publishing and open access data in scientific research.

Related issues such as a classification for project datasets and an overview of copyright licensing for open access are also discussed. Each project dataset is then described in detail in Section 3, using a standardised template.

4. APPROACH TO DATA MANAGEMENT

1. Open access

Open access (OA) is understood as the free, online provision of re-useable scientific information to other users. There are many good reasons to make the data and findings from publicly funded research openly available to the research community, the commercial sector and civil society.

Data availability is therefore categorised at this stage in one of three ways:

- Open Data that is shared for re-use or that underpins a scientific publication.
- Consortium Confidential data that is accessible to all consortium partners and the Local Implementation Groups (under certain specific conditions described in the project Consortium Agreement) but retained within the consortium.
- Private Data that is maintained by an individual partner for his own purposes.

Project datasets for dissemination will be open access by default, at the very least to validate scientific publications. However, not all of the project work packages will produce datasets that are intended for public dissemination; much of the data created and stored during the project is for internal management and communication within the consortium only. It is therefore likely that much of the data is categorised as Consortium level data. However, the project will seek to openly disseminate its research findings, except in cases where there are defined exploitable outcomes, privacy concerns or there will be a high administrative burden for a dataset or limited worth to other users. The two main aspects of this dissemination approach are open access to scientific publications and open access to research data. Each is considered in the following sections.

2. Open data

There are four main aspects of open data summarised in the acronym FAIR:

- Findable- data has a unique, persistent ID, located in a searchable resource, and documented with meaningful metadata.
- Accessible- data is readily and freely retrievable using common methods and protocols, metadata is accessible even if the data is not.
- Interoperable- data is presented in broadly recognised standard formats, vocabularies and languages.
- Re-useable- data has clear licences, and accurate meaningful metadata conforming to relevant community standards and identifying its content and provenance.

The data management plan establishes how this approach will be realised in practice with the initial plan presenting an overview and greater detail provided in the interim and final reports as the work packages proceed.

Project datasets for dissemination will be open access by default, at the very least to validate scientific publications. However, not all of the project work packages will produce datasets that are intended for public dissemination; much of the data created and stored during the project is for internal management and communication within the consortium only. The distinction between datasets is clearly described in the dataset templates.

3. Data storage and sharing

The project has five main data storage and sharing facilities according to the type of data and its intended accessibility.

- Private data: Stored locally on partner storage systems, subject to institutional back up practices.
- Consortium data: Uploaded to the project collaboration platform Microsoft Teams, a cloud storage system for simple and secure access for all partners from a web browser. Data is maintained with regular offsite backups.
- Open data: Three facilities will be used during the project:
 - The project website, managed by R2M, will be the first point of contact for public dissemination. It will host project technical reports and other materials such as events listings, blog articles, images, videos, links to partner organisations and related projects.
 - Large, re-useable data sets will be deposited in an open data repository [e.g. Zenodo] selected by the task leaders during the delivery of the relevant work packages.
 - OpenAIRE, to which all datasets related to the project will be linked.

4. Copyright licences

When material is widely shared, copyright licences protect the authors of work and grant specific rights to publishers and others to use this work. The European Commission encourages authors to retain their copyright whilst disseminating it as open access. Creative Commons (CC, ref 5) provides legal tools to enable open access in these circumstances, with CC-BY (licensees may copy, distribute, display and perform the work and make derivative works and remixes based on it only if they give the author or licensor the credits in the manner specified by these or by any reasonable other manner) and CC0 (freeing content globally without restrictions) enabling re-use by third parties.

Where research findings are published in a journal or other scientific outlet, there should be consideration of the copyright agreement with the publishers, which may involve an embargo period. Submission in the REGEN repository requires the author to agree to a non-exclusive distribution licence, and a Creative Commons licence may be added at this stage.

At this initial stage, it is not possible to define the copyright arrangement for each project dataset. The most appropriate licensing arrangements for each of the project datasets will be investigated as they are better characterised by their respective work packages. The final data management plan will be updated to that effect.

5. DESCRIPTION OF PROJECT DATASETS

A dataset is a collection of related data, collected or created, including any existing data or third-party sources, which will be used or reused during and beyond the lifetime of the REGEN project. The description of the datasets below is based upon the Horizon Europe DMP template.

The template is a set of questions that each Data Manager responsible for a specific dataset should answer with a level of detail appropriate to the project. The DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur.

1. List of collected datasets

This first version of the Data Management Plan mostly addresses the most important data that will be handled in the REGEN project. A list of the most important datasets that have been identified as of June 2024 in REGEN is available below. It should be noted that this list is likely to evolve depending on the evolution of the various activities foreseen in the project.

The datasets are numbered according to their primary work package and task number, as laid out in the REGEN DoA.

Dataset name	Data manager
T1.3.1 Market Gap Analysis	LISER
T1.3.4 Training program	R2M
T2.3.1 Technology Component (TC5) : DinyCont Flows	DINYCON
T2.3.2 Technology Component (TC4) : ReSuite	DEMO

T2.3.3 Technology Component (TC2) : ICL Platform	IES
T2.3.4 Technology Component (TC1) : MUST	LIST
T2.3.5 Technology Component (TC3) : Digital Deconstruction Platform	LIST
T2.3.6 Technology Component (TC6) : Nbenefit\$	LIST
T2.3.8 Technology Component (TC) : Map Server	LIST
T2.4 Survey on decarbonised behaviours	LISER
T3.1 Demo sites monitoring	TECNALIA
T3.2 Demo site Beckerich	Beckerich municipality
T3.3 Demo site Laredo	Laredo municipality
T3.4 Demo site Milan	Milan municipality
T3.5 Demo site Dublin	Dublin City Council
T3.6 Evaluation of demo sites	DEMO
T4.1 Impact evaluation dataset	NOBATEK
T4.3 AFUR replication assessment via 10 virtual cases	ISOCARP
T4.4 Application of AFUR on 30 additional interventions	NOBATEK
T4.5 Advocacy material and roadmap	DEMO

2. The form - template

Information about each dataset has been collated by Task Leaders in the format presented below.

Number and name of the dataset - template	
1. DATA SUMMARY	
Task & Data Manager	Work Package and Task numbers related to the dataset, and the Data Manager who takes responsibility
Availability	Private, Consortium or Open, as defined in section 2.3
Purpose of the data	What is the purpose of the data collection/generation and its relation to the objectives of the project?
Origin of the data	What is the origin of the data? Produced in the framework of the project or reused?

Storage	Where will the data be stored? (e.g: open data repository like Zenodo, REGEN website for public deliverables, Microsoft Teams) as defined in section 2.3
Data nature and size	What types and formats of data will the project generate/collect? What is the expected size of the data?
Data utility	To whom might your data be useful, at Consortium level or external to the project?
Dataset Specific Metadata	Keyword(s) that categorize data to make it linked/searchable
2. FAIR DATA	
Making data findable	Will the data be accessible through a free and standardized access protocol? What information is needed for the data to be read and interpreted in the future? how this documentation and metadata will be captured / created? what metadata standards will be used and why?
Making data openly accessible	Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.
Making data interoperable	In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?
Increase data reuse	How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)? Will the provenance of the data be thoroughly documented using the appropriate standards?

3. ALLOCATION OF RESOURCES	
Costs	What are the costs for making data FAIR in your project? How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions).
Long term preservation	Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?
4. DATA SECURITY	
Data security, recovery and storage	What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? Is the data safely stored in certified repositories for long term preservation and curation?
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. How will you protect the identity of participants if required? e.g. via anonymisation
Informed consent	Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?
6. OTHER ISSUES	
Other national/funder/sectorial/departamental procedures	Do you make use of other national/funder/sectorial/departamental procedures for data management? If yes, which ones?

3. Data collection in WP1

T1.3.1 Capacity Building needs assessment, REGEN

1. DATA SUMMARY

Task & Data Manager	Task 1.3: Capacity Building needs assessment and market gap analysis <u>Subtask1.3.1</u>: Capacity Building needs assessment <u>Manager</u>: Prof. Martin DIJST, Luxembourg Institute of Socio-Economic Research (LISER)
Availability	Open Science because based on open cooperative work (co-creation), tools (participatory methods) and science promotion (citizen science).
Purpose of the data	The aim of this subtask is to <i>"define the key competencies and skills needed to develop a training program for 'coaches'."</i> The capacity-building plan will be developed by LISER, together with input from other partners involved in the demo-sites. • After an initial roundtable clarifying baselines and main barriers specific to each group, structured interviews (including stakeholder surveys) will be conducted to assess the needs in both technical and 'soft' skills that would support the regeneration of the neighborhoods. • Participatory meetings (in-presence meetings and online meetings) will complete the collection of competencies and skills needed. • Demo-site leaders will translate and roll this out at each demo-site and compile results into a report on training needs for each stakeholder group identified as key.
Origin of the data	The input on the competencies and skills needs for the 'coaches' will be collected from the partners involved in the demo-sites mainly through semi-structured interviews (complemented with a short survey) and participatory meetings (in presence and/or online).
Storage	REGEN website for public deliverables, Microsoft Teams) as defined in section 2.3. All project data will be stored and secured in the LISER Research Data Infrastructure (RDI).
Data nature and size	Semi-structured interviews (12-16 interviews): qualitative data (audio data, anonymized transcriptions) Participatory meetings (4-6 meetings): qualitative data (audio/video data, anonymized transcriptions)
Data utility	The data collected during this subtask will be used to define the competences and skills needs for the coaches.

	<u>At LISER level</u>, the data collected is useful for the researchers and partners involved to define the profile of the coaches. The aim is to have coaches able to empower (capacity building) other citizens. <u>At the consortium level</u>, the data will help in training coaches as part of urban regeneration processes in the demo sites selected for the project (Beckerich, Laredo, Milano, Dublin). The coaches will be trained to Assessment Framework for Urban Regeneration and act as engagers towards other citizens and the data will contribute to the development AFUR. <u>At the external EU level</u>, the data is of potential interest for researchers involved in similar topics / in similar methodologies.
Dataset Specific Metadata	Capacity building, empowerment, co-creation, sustainable learning
2. FAIR DATA	
Making data findable	Each dataset will be labelled with detailed metadata, covering aspects such as data type (here: mainly qualitative), collection method (audio, video, observation notes), methodology (semi-structured interviews/ participatory meeting) and geographic location to improve discoverability.
Making data openly accessible	While primary datasets involving personal information will be under restricted access, aggregate data and findings will be made available through open access platforms, respecting GDPR and ethical guidelines.
Making data interoperable	We will utilize standard data formats and widely accepted metadata standards to ensure interoperability and facilitate seamless integration with other datasets.
Increase data reuse	A detailed documentation about the interview guides, the coding and categorizations, the matrix, will be available to facilitate the reuse.
3. ALLOCATION OF RESOURCES	
Costs	The costs associated with data management, primarily human resources expenses, are covered by LISER's budget. Infrastructural costs are covered by LISER's indirect costs (overhead).
Long term preservation	The project REGEN has 48 months (01/01/2024 – 31/12/2027). For publication purposes, replication purposes, and potential spillover, pseudo anonymized data will be

	stored and available five more years following the end of the project (2032).
4. DATA SECURITY	
Data security, recovery and storage	<i>Storage</i> During the data analysis phase will be performed on LISER servers (in Luxembourg at LISER premises). - Raw data, intermediate analysis data and final data will be stored in a secure space of the LISER Research Data Infrastructure (dedicated server "RDI"). They will all have a special level of access dedicated to each of them (either restricted and secure). <i>Security</i> - Access to the server, protected by a firewall, requires identification and (strong) authentication by authorized users only, i.e. the managers/administrators of the LISER research databases ; - Identification passwords are subject to an internal policy defining both the format (combination of at least 8 alphanumeric and special characters) and also a regular frequency of password renewal; - Physical access to the server can be achieved by dedicated technical personnel (system administrator) ensuring the security and integrity of the infrastructure. <i>Recovery</i> - For each collected data/measurement, a codebook will be generated, and codes and syntaxes will be save at every step. Data of different version will be successively store to provide the ability to identify up-to-date version, last addition and limit the risk of data loss. - All actions are logged in logs stored on the archiving platform; - Each action (deposit, extraction, etc.) will give rise to a recording in an access register; - Archived data files are encrypted, the encryption keys of which are known only to the LISER research database managers.
5. ETHICAL ASPECTS	

Ethical and legal issues on data sharing	All participants will be assigned code numbers (numeric identifiers) via a pseudonymization key list. After this step all digital data will be labelled with these code numbers and do not contain any personal identifiers. The pseudonymisation key list linking the code numbers to personal data of participants (e.g. email addresses) will be securely stored on the LISER server. The list will be stored separately from any raw or processed data. After a retention period of 5 years, the key list as well as the raw and processed data will be destroyed. The European General Data Protection Regulation will be respected at all times.
Informed consent	All participants have to sign a consent form, providing them with information about the use, storage, and rights concerning their data.
6. OTHER ISSUES	
Other national/funder/sectorial/departamental procedures	Complies with Luxembourg's national data protection laws in line with GDPR, Horizon Europe requirements, and specifically the mandate to make data collected through the REGEN process OPEN ACCESS.

T1.3.2 Training program

1. DATA SUMMARY	
Task & Data Manager	Work Package: 1 Task number: T1.3 (Subtask 1.3.4) Data Manager: R2MF
Availability	Open/Public
Purpose of the data	The purpose of data collection and generation is multifaceted and aligns closely with REGEN's overarching objectives. Here's a breakdown of how this data supports the training program and the broader goals of neighborhood regeneration: • Empowering Local Communities • Enhancing Stakeholder Collaboration • Building Sustainable Skills • Promoting Inclusive Development • Improving Project Implementation and Outcomes
Origin of the data	This subtask (T1.3.4 – lead: R2M, participants: all partners) starts at M24 and the origin of the collected data will be determined upon its kick-off.

Storage	REGEN website, public deliverable (D1.4) and Microsoft Teams for internal purposes and access of project partners.
Data nature and size	The training program will generate/collect participant demographics, session attendance, feedback surveys, training materials (slides / handouts), and audio-visual recordings. Formats include CSV, PDF, and MP4. The expected size is modest, typically under 500 MB, considering the limited number of sessions and the nature of the data.
Data utility	The data will be valuable to the demosite leaders (ie. Municipalities) and other consortium members for evaluating training effectiveness and refining future programs. Externally, it benefits local authorities, urban planners, and community organisations seeking insights into best practices and methodologies for neighborhood regeneration and capacity building initiatives.
Dataset Specific Metadata	"Capacity Building, Neighborhood Regeneration, Training Materials, Participant Feedback, Attendance Data"
2. FAIR DATA	
Making data findable	Yes, data will be accessible via Open Access modes. Documentation will include detailed descriptions, collection methods, and context. Metadata will follow Dublin Core standards for interoperability and ease of future interpretation.
Making data openly accessible	Yes, All data will be openly available on REGEN website, public deliverable (D1.4) and Microsoft Teams for internal purposes and access of project partners.
Making data interoperable	N/A
Increase data reuse	This subtask (T1.3.4 – lead: R2M, participants: all partners) starts at M24 and the origin of the collected data will be determined upon its kick-off.
3. ALLOCATION OF RESOURCES	

Costs	This subtask (T1.3.4 – lead: R2M, participants: all partners) starts at M24 and the origin of the collected data will be determined upon its kick-off.
Long term preservation	/
4. DATA SECURITY	
Data security, recovery and storage	This subtask (T1.3.4 – lead: R2M, participants: all partners) starts at M24 and the origin of the collected data will be determined upon its kick-off.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Ethical and legal issues include participant privacy and data protection. Identity will be protected through anonymization techniques and compliance with GDPR, ensuring sensitive data is securely handled and shared responsibly.
Informed consent	Yes, the future questionnaires formed in this task (T1.3.4) will include informed consent for data sharing and long-term preservation, explicitly detailing participant rights, data usage, and retention policies to comply with GDPR and ethical standards.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	This subtask (T1.3.4 – lead: R2M, participants: all partners) starts at M24 and the origin of the collected data will be determined upon its kick-off.

4. Data collection in WP2

T2.3.1 Technology Component (TC5) : DinyCont Flows	
1. DATA SUMMARY	
Task & Data Manager	WP3: Task 3.3, Task 3.4. Manager: Pablo Candelas.
Availability	Open.

Purpose of the data	The purpose is the analysis of the mobility patterns and the accessibility in the Laredo and Milan demo sites. This will help characterize the locations and it's needs and help determine the effectiveness of the interventions.
Origin of the data	The data is produced in the framework of the project through the installation of sensors in the demo sites.
Storage	The data will be stored in the private servers of Dinycon, as well as in an open data repository.
Data nature and size	Historic counting reports in csv, excel or pdf formats. Expected size: ~1000 KB.
Data utility	Useful at Consortium level for Demo site administrators and researchers.
Dataset Specific Metadata	People/vehicle, counting, report, <i>dates</i> , Laredo, Milan.
2. FAIR DATA	
Making data findable	The data will be accessible through the standard protocols agreed on by the partners upon setting up the open data repository.
Making data openly accessible	All data will be openly available through the open data repository.
Making data interoperable	In case we must use uncommon or generate project specific ontologies or vocabularies, we will provide mappings to more commonly used ontologies.
Increase data reuse	/
3. ALLOCATION OF RESOURCES	
Costs	Personnel costs for the time needed to make data FAIR, covered in the Agreement.
Long term preservation	The data will be kept in the open repository for as long as it remains active. It will also be preserved in Dinycon's private servers indefinitely.
4. DATA SECURITY	

Data security, recovery and storage	The data will be stored in the private servers of Dinycon, as well as in an open repository. No sensitive data.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	No ethical or legal issues.
Informed consent	No personal data.
6. OTHER ISSUES	
Other national/funder/sectorial/departamental procedures	/

Task 2.3.2 - Technology Component (TC4) : ReSuite	
1. DATA SUMMARY	
Task & Data Manager	WP2, Task 2.3.2 - Technological developments implementation Task 2.3.2 is led by IES, with contributions from LIST and all technical partners. DEMO is responsible for TC4 (RE Suite)
Availability	Only available for the consortium members.
Purpose of the data	Task 2.3.2 aims to facilitate the integration of various digital tools within the REGEN project. This includes the integration of the IES ICL tools with the LIST MUST software system, developing APIs for data exchange and ensuring interoperability with other tools in the REGEN toolbox. The purpose of collected data by RE Suite is to assess the condition of the assets and optimize the maintenance plans.
Origin of the data	Produced within the framework of the REGEN project or data available in the other data bases.
Storage	Typically RE Suite data is stored in a private server in the Azure Cloud.
Data nature and size	The data types include software code, APIs, integration protocols and possibly user interaction data from the digital tools. Expected formats include JSON, XML, JPEG, vector data and CSV for data exchange

	formats and source code files in languages such as Python, C#, or others used in software development. The size of the data is yet to be determined and will depend on the volume of interactions and integrations.
Data utility	Useful for technical partners and stakeholders involved in the REGEN project for monitoring and optimizing urban regeneration interventions.
Dataset Specific Metadata	At the moment we are exploring data. None
2. FAIR DATA	
Making data findable	Data will be accessible through standardized API protocols. Documentation and metadata will be captured using established metadata standards for interoperability, such as the demo site Dublin.
Making data openly accessible	Not all data will be available openly due to privacy and security concerns. Certain datasets will be restricted to the consortium, as specified in the Grant Agreement.
Making data interoperable	The dataset will use common ontologies and provide mappings to widely used standards to ensure interoperability.
Increase data reuse	If needed, data documentation will be provided, including methodology, codebooks, and data-cleaning procedures to facilitate reuse.
3. ALLOCATION OF RESOURCES	
Costs	Costs for making data FAIR are covered by the Horizon Europe grant.
Long term preservation	The data management plan has yet to determine what data to preserve and for how long.
4. DATA SECURITY	
Data security, recovery and storage	Data is stored securely in certified repositories with provisions for data recovery and secure transfer. Data security measures will be in place, including secure storage, data recovery plans, and transfer protocols for sensitive data. After discussion with other

	partners, certified repositories will be used for long-term preservation and curation.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Ethical considerations will be addressed in accordance with the approved ethics review and relevant guidelines. To ensure privacy and confidentiality, anonymization techniques can be implemented.
Informed consent	Informed consent will be obtained for any personal data collected, ensuring compliance with GDPR and other relevant regulations.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	No

T2.3.3 Technology Component (TC2) : ICL Platform	
1. DATA SUMMARY	
Task & Data Manager	WP2 Task 2.2: Definition of a software architecture for REGEN Task 2.3: Technological deployment
Availability	Data relating to results will be open i.e. CSV, GeoJSON generated for REGEN. Tools relating to IES ICL are protected IP and are not freely available to the public. Data relating to IoT sensor data are generally open; however, their continued access depends on the relative project partners commitments to the GA
Purpose of the data	Data consumed by the ICL tool will be used to inform context towards accurate simulations such as building information, socio-economic data, infrastructure data
Origin of the data	Generally, data is derived from local sites from government, NGO, university, or private resources
Storage	Data is often updated to Zenodo at the end of project lifespan; however, alternative data repositories including REGEN website is also applicable

Data nature and size	GeoJSON, CSV, DWG, SHP, JSON, 1 -15 GBs
Data utility	Consortium level as well as external local government, universities, and interested enterprises both for profit and non-profit
Dataset Specific Metadata	Building Footprint data, Power Network Data, Energy Performance Certificate Data, Environmental Data, Local Weather Files.
2. FAIR DATA	
Making data findable	Data that is earmarked to be open, data labeling will follow standard themes associated with similar data bespoke to each filed i.e. planning and environmental policy naming convention for large GIS building datasets as “Building footprint data”
Making data openly accessible	Data will only be released under the instruction of the original data author and the project coordinators based on instructions from the original data author. If data contains GDPR sensitive information or has monitory IP important to the data author, they may request for data to be private or embargoed for a certain amount of time.
Making data interoperable	Ontologies are standards for the ICL. It's suggested that an ontology should be derived for REGEN project or a standard ontology be applied based of ISO standards. Data formats are critical and we do not accept certain formats
Increase data reuse	Documentation is not the responsibility of IES towards using generated data in other platforms. This will be up to the related stakeholders who are promoting the results form our tools towards the goals they have set out related to their WP or tasks
3. ALLOCATION OF RESOURCES	
Costs	Potential costs are related data storage; however, there are no costs envisioned for data storage in the REGEN project related to IES
Long term preservation	If IES tools are accessed by REGEN project partners after the REGEN project ends, the user will have to pay for a license to access the ICL

4. DATA SECURITY	
Data security, recovery and storage	Data stored on IES servers are protected under third party contracts with our web service provider. These contracts protect cloud-based data in the events of server outage as well as high standards of encryption for security issues.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Data with sensitive personal information (GDPR) will need to be proceeds appropriately by relevant patterns before it is shared with IES. IES anonymizes and pseudonymized data before processing.
Informed consent	Informed consent happens before data is received by IES.
6. OTHER ISSUES	
Other national/funder/sectorial/de partmental procedures	Ensure appropriate data format shared by partners. I.e. GeoJSON or SHP only for iCD master planning simulations. We only except certain formats, we do not convert existing data set.

T2.3.4 Technology Component (TC1) : MUST	
1. DATA SUMMARY	
Task & Data Manager	WP2. T2.3
Availability	Consortium
Purpose of the data	Data collected by MUST aims at gauging the priorities and preferences for a given set of criteria for the audience participating in the workshops. Collected data will guide the process of deploying the interventions.
Origin of the data	Data collected by MUST is generated during multiple workshops held on the various pilot sites.
Storage	Collected data will be stored locally (pilot site) in a dedicated database and extracted for statistical evaluation at a later stage.

Data nature and size	Collected data mostly consists of numeric rankings for pilot site specific criteria. Spatial organigram of the site (if applicable) will be stored as geometries in the database. Data size depends on the number of workshops/sessions, however, knowing that the size of a workshop/session dataset is relatively small ($\ll$ 1Mbyte), we're not expecting huge amounts of data.
Data utility	Collected data will guide pilot site managers in selecting the most appropriate interventions.
Dataset Specific Metadata	/
2. FAIR DATA	
Making data findable	Not applicable
Making data openly accessible	Not applicable
Making data interoperable	Not applicable
Increase data reuse	Not applicable
3. ALLOCATION OF RESOURCES	
Costs	Not applicable
Long term preservation	Not applicable
4. DATA SECURITY	
Data security, recovery and storage	Collected raw data will be extracted and statistically evaluated. Raw data may be disposed of after this stage.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Collected raw data represents the (anonymous) consensus of a group of workshop/session participants. Collected data does not contain or refer to any personal data of participants. No personal data will be collected.

Informed consent	No personal data will be collected. Not applicable
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	NA

T2.3.5 Technology Component (TC3) : Digital Deconstruction Platform

1. DATA SUMMARY	
Task & Data Manager	WP2 Annie Guerriero
Availability	Consortium (maybe a subset will be made open later on)
Purpose of the data	This dataset reflects all the data gathered for the application of Digital Deconstruction process on several buildings in Milan and Beckerich. This is foreseen to contribute to the achievement of a.o. REGEN Expected Outcome 3.
Origin of the data	The data is produced within the project by the partners (or subcontractors), in relation with demo sites activities of WP3.
Storage	Data is stored within LIST IT systems.
Data nature and size	The data includes 3D scans, BIM models, reuse potential indexes (RBIM) and material passports and impact calculations. In addition the platform collects usernames and email addresses.
Data utility	Data will be used by selected consortium partners to visualize data useful for Milan and Beckerich demo sites.
Dataset Specific Metadata	BIM IFC, Uniclass, NL-SFB
2. FAIR DATA	

Making data findable	Data are assigned a globally unique and persistent identifier
Making data openly accessible	Not all data will be available openly due to privacy and security concerns. Data will be accessible through standardized API protocols.
Making data interoperable	The dataset will use common ontologies and provide mappings to widely used standards, incl. classifications, to ensure interoperability.
Increase data reuse	Data usage license is defined following the terms of the Consortium Agreement.
3. ALLOCATION OF RESOURCES	
Costs	NC
Long term preservation	NC
4. DATA SECURITY	
Data security, recovery and storage	Data will be stored in LIST servers.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	NA
Informed consent	NA
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	NA

T2.3.6 Technology Component (TC6) : Nbenefit\$

1. DATA SUMMARY	
Task & Data Manager	Work Package and Task numbers related to the dataset, and the Data Manager who takes responsibility
Availability	Input data collected to generate the archetype will be public.

	Broadleaves and conifers archetypes (general archetypes) will be public. The aggregated indicators of environmental costs and benefits resulting from the use of NBenefit\$ for the study areas on Milan and Laredo will be public. Specific tree archetypes for the tree species used in the study areas in Milan and Laredo will be shared with those partners within the consortium to allow them to test the NBenefit\$ tool). The larger set of tree archetypes (species specific) simulated for the cities of Laredo and Milan will be private since the work to generate them will be funded via another project (FNR JUMP22 / 17746254 / NBenefit\$ 3.0).
Purpose of the data	Cost benefit analysis of tree nature-based solutions.
Origin of the data	The input data to generate the archetypes will be collected by the project. The modelling framework to generate the archetype is reused from previous projects. The new archetypes for Milan and Laredo will be generated by another project and reused in REGEN.
Storage	Input data: MS Teams Archetypes: LIST Server
Data nature and size	All data will be csv data. The size depends on the number of tree species to be generated to replicate the conditions in the intervention areas.
Data utility	Urban planner and citizens in Milan and Laredo
Dataset Specific Metadata	archetypes, nature based solution, trees, cost benefit analysis
2. FAIR DATA	
Making data findable	/

Making data openly accessible	Trees archetype will be private because the generation of these archetypes is funded by another project. In REGEN we will use these archetypes to help the cities to support the implementation of NBS projects, but only the final outcome of the calculations will be shared with the partners.
Making data interoperable	/
Increase data reuse	Readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement and data provenance will be generated for each archetype.
3. ALLOCATION OF RESOURCES	
Costs	/
Long term preservation	Aligned to the project protocol
4. DATA SECURITY	
Data security, recovery and storage	Data will be stored in LIST servers.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	No personal data will be collected
Informed consent	No personal data will be collected
6. OTHER ISSUES	
Other national/funder/sectorial/departamental procedures	The archetypes area generated using the FNR JUMP funding. The aim of the project is to valorize the decision support tool NBenefit\$. For this reason, the tree archetypes which are at the core of the tool, cannot be make publicly available.

T2.3.8 Technology Component (TC) : Map Server

1. DATA SUMMARY	
Task & Data Manager	WP2 REGEN data collection on Mapserver Philippe PINHEIRO
Availability	Open and Private

Purpose of the data	Visualize GIS data, compute models based on GIS data, create statistics and dashboards
Origin of the data	Public open data + Municipalities public and private data
Storage	LIST servers
Data nature and size	Geotiff, ESRI Shapefile, GeoJSON (Tb of data)
Data utility	Municipalities population and public
Dataset Specific Metadata	GIS
2. FAIR DATA	
Making data findable	Yes, using OGC Standards related to GIS like WMS, WFS, WPS, Maps, etc.
Making data openly accessible	No, not all data will be made openly available because some data will have restricted access conditions
Making data interoperable	We will only have interoperable data
Increase data reuse	With OGC standards metadata describing the datasets Will the provenance of the data be thoroughly documented using the appropriate standards? Yes, using OGC standards metadata
3. ALLOCATION OF RESOURCES	
Costs	All costs are covered by the project budget or LIST central IT services. LIST will do its best to provide long term access to these datasets as part of his mission as a public Research and Technology Organization
Long term preservation	Municipalities and LIST will decide whether the dataset should be maintained or not. Indeed, it might be useful to the municipalities to demonstrate the impacts of REGEN interventions over time.
4. DATA SECURITY	

Data security, recovery and storage	LIST security team provisions Is the data safely stored in certified repositories for long term preservation and curation? Yes
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	All data will be anonymized
Informed consent	No personal data is used
6. OTHER ISSUES	
Other national/funder/sectorial/departamental procedures	No

T2.4 Survey on Decarbonised Behaviours

1. DATA SUMMARY	
Task & Data Manager	Task 2.4 Survey on Decarbonised Behaviours Manager: Prof. Martin DIJST, Luxembourg Institute of Socio-Economic Research (LISER)
Availability	Open Access. However, the qualitative dataset will be anonymized and transcripts made publicly available. Datasets about behavioral patterns such as mobility trajectories, consumption patterns and individual aspirations will be pseudonimized. This is because personal daily habits of the households and spatial components of the data constitute identifying information so the data will not be shared openly. Based on request and subsequent evaluation form the LISER Data Center, data provision contract mentioning among other points, rights, duties, disclaimers, data property and usage, and publication rights will be considered.
Purpose of the data	Aim: Collect data regarding to sustainable behaviours and lifestyles of people and opportunities to change these to reduce their carbon footprint in terms of energy consumption and mobility behaviour This data relates to the understanding and development of sustainable behaviours (with respect to energy use/consumption and activity/travel behaviour) to build

	knowledge about sustainable solutions regards to climate emergency. As a part of REGEN project, the survey will be deployed in Beckerich first, then in different cities and neighbourhoods (La Puebla Vieja de Laredo in Spain, Giambellino-Lorenteggio in Milan in Italy and the Irishtown & Ringsend area in Dublin in Ireland) with the support of the local authorities.
Origin of the data	The data will be collected by filling out an online questionnaire that constitutes the survey. This questionnaire is composed of four parts: Part 1: General sociodemographic information This part collects general sociodemographic data: e.g., age, sex, marital status, highest educational degree, occupation, financial situation Part 2: Revealed preferences (one subpart for energy equipment, energy consumption and habits, a second subpart for transports, daily activities and travel behaviour) This part of the survey investigates: • Respondent's' energy consumption (energy performance certificate, sources, cost, equipment, habits) • Respondents' typical trip and habits (in terms of equipment, activities, choice of transport mode, frequency use and cost) Part 3: Stated preferences (one subpart for energy equipment, energy consumption and habits, a second subpart for transports and daily trips) This part of the survey investigates: • Stated preference experiments are tested under which scenarios people are willing to change energy source, to reduce energy use, and to buy energy efficient building technical systems. • Stated preference experiments will be designed to measure under which condition further substitution of fossil-fueled car trips by sustainable alternatives and offline by online activities will be chosen. Also, the buying of electric cars will be included in scenarios. Part 4: Opinions about environmental concerns and changes This part of the survey investigates

	• Environmental issues the respondent is concerned with • Environmental impact • Risk perception LISER is responsible for setting out the principles of the data collection, the design of an online survey, and the data analysis. Local authorities will support LISER in collecting the requested data in their respective city.
Storage	REGEN website for public deliverables, Microsoft Teams) as defined in section 2.3. Beyond that, all project data will be stored and secured in the LISER Research Data Infrastructure (RDI). Thus, the raw data from Qualtrics, the resulting final data, but also the associated metadata, the programs and all technical and methodological documentation deemed useful and necessary will be saved in compliance with the legal conditions of conservation. This data will be stored at LISER and documented with a codebook
Data nature and size	Raw data: True responses from the respondents to the questionnaire Pseudo anonymized data: Data where direct personal identifiers (name, email address) about respondents are removed to protect and safeguard their privacy and confidentiality Fully anonymized data: Data where all direct and indirect personal identifiers are removed (e.g: exact home address replaced by district) and/or aggregated (e.g: age converted into age categories). Overall, the mixed-type data (quantitative from survey selections and qualitative from open-ended responses), is expected to involve approximately 20% of the 14000 households situated in the four demo sites of the project. 200 households respondents from Beckerich and Laredo, 2000 for Milano and 400 for Dublin Current <u>Research design</u> : The current study represents an observational cross-sectional study. <u>Methods of recruitment and data collection</u> : Recruitment will be done through partners and the local authorities of each case study.

		A specific person hired by the respective local authorities will inform people living in each case study about the survey. <u>Statistical analysis:</u> The statistical analysis (descriptive statistics, multivariate analyses, GIS-analysis) will be performed using standard software such as STATA, SPSS, R, Python.
Data utility		At an internal level, the data will help researchers from LISER understand behavioral patterns related to urban regeneration processes. At the consortium level, the data will help in developing people centered interventions as part of urban regeneration processes in the demo sites selected for the project. Beyond this, the data will contribute to the development of an innovative Assessment Framework for Urban Regeneration (AFUR). At the external EU level, the data can be of interest to researchers and local practitioners studying/interested in urban sustainability and behavior change.
Dataset Metadata	Specific	Decarbonisation behaviors, urban sustainability, energy consumption, sustainable mobility, environmental attitudes.
2. FAIR DATA		
Making data findable		All datasets will be cataloged with comprehensive metadata including data type (qualitative/quantitative), collection method (questionnaire/survey), and geographic location to enhance discoverability.
Making data openly accessible		While primary datasets involving personal information will be under restricted access, aggregate data and findings will be made available through open access platforms, respecting GDPR and ethical guidelines.
Making data interoperable	data	We will use standard data formats (e.g., CSV, SPSS) and common metadata standards to ensure interoperability and ease of integration with other datasets.

Increase data reuse	Detailed documentation including codebooks, README files, and methodological descriptions will accompany the dataset to facilitate reuse. Provenance and change logs will be maintained to document dataset evolution.
3. ALLOCATION OF RESOURCES	
Costs	Estimated costs for data management (human resources) are accounted for through LISER's budget. Infrastructural costs are covered by LISER's indirect costs (overhead).
Long term preservation	The project REGEN has 48 months (1 st January 2024 – 31 st December 2027). For publication purposes, replication purposes, and potential spillover, pseudo anonymized data will be stored and available five more years following the end of the project (2032).
4. DATA SECURITY	
Data security, recovery and storage	Storage During the data collection phase, the data is stored on data centers located in Germany by Amazon web services (under contract with Qualtrics). During the data analysis phase will be performed on LISER servers (in Luxembourg at LISER premises). - Raw data, intermediate analysis data and final data will be stored in a secure space of the LISER Research Data Infrastructure (dedicated server "RDI"). They will all have a special level of access dedicated to each of them (either restricted and secure). Security - Access to the server, protected by a firewall, requires identification and (strong) authentication by authorized users only, i.e. the managers/administrators of the LISER research databases ; - Identification passwords are subject to an internal policy defining both the format (combination of at least 8 alphanumeric and special characters) and also a regular frequency of password renewal;

	- Physical access to the server can be achieved by dedicated technical personnel (system administrator) ensuring the security and integrity of the infrastructure. <i>Recovery</i> - For each collected data/measurement, a codebook will be generated, and codes and syntaxes will be save at every step. Data of different version will be successively store to provide the ability to identify up-to-date version, last addition and limit the risk of data loss. - All actions are logged in logs stored on the archiving platform; - Each action (deposit, extraction, etc.) will give rise to a recording in an access register; - Archived data files are encrypted, the encryption keys of which are known only to the LISER research database managers.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	All participants will be assigned code numbers (numeric identifiers) via a pseudonymization key list. After this step all digital data will be labelled with these code numbers and do not contain any personal identifiers. The pseudonymisation key list linking the code numbers to personal data of participants (e.g. email addresses) will be securely stored on the LISER server. The list will be stored separately from any raw or processed data. After a retention period of 5 years, the key list as well as the raw and processed data will be destroyed. LISER ISP will unlink the responses from the residential location information. They will put the residential location information in a separate file with an ID number. This file will be sent to a third person (a researcher who is not very closely involved in the analysis). This researcher will then geo-code the information and add spatial variables to it. They will send this data back to the ISP. The ISP will then merge this geo-located information with the responses of the questionnaire. This new file will be sent to the researcher(s) working on the project. This way, the researchers will get the aggregated social and spatial information, but in a pseudonymized way. The European General Data Protection Regulation will be respected at all times.

	We have already sought and obtained ethical approval from the Ethics Review Board of LISER who has been provided all details related to the data collection process.
Informed consent	Consent will be explicitly obtained from all participants, detailing the use, storage, and rights concerning their data. This will be done electronically before they begin answering the questionnaires. Participants will also have the choice to be informed about the results of the research process.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	Adheres to Luxembourg's national data protection laws in accordance to GDPR, Horizon Europe requirements, and more specifically, requirements to make data gathered as part of the REGEN process OPEN ACCESS.

5. Data collection in WP3

T3.1. Demo site data for AFUR evaluation	
1. DATA SUMMARY	
Task & Data Manager	WP3, task 3.1 Demosite data collection leader: Beckerich – LIST (Sylvain Kubicki temp) Laredo – TEC (Elena Usobiaga) Milan – AMAT (Andrea Naldini) Dublin – DCC (Darby Mullen)
Availability	Consortium
Purpose of the data	Apply the detailed monitoring scheme developed in Task 2.5 to evaluate the demo site impacts via AFUR methodology
Origin of the data	Mainly obtained from existing sources.
Storage	REGEN Sharepoint
Data nature and size	Pdf, Xls, csv, ... Expected size: not defined yet

Data utility	Consortium level (for demos monitoring)
Dataset Specific Metadata	Data category (e.g. energy, mobility, air pollution...), location, time representativeness, data generator, data source
2. FAIR DATA	
Making data findable	documents (.pdf), tables (.xls, .csv), optionally semantic graph data (.xml)
Making data openly accessible	NO, only the original outputs of REGEN. source data will be referenced where possible. References to alternative sources of data will be provided.
Making data interoperable	Not defined yet
Increase data reuse	Not defined yet
3. ALLOCATION OF RESOURCES	
Costs	Not defined yet
Long term preservation	Not defined yet
4. DATA SECURITY	
Data security, recovery and storage	Not defined yet
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	No personal data will be expected and if necessary, it will be previously anonymized.
Informed consent	N/A
6. OTHER ISSUES	
Other national/funder/sectorial/departamental procedures	Not defined yet

T3.2 Demo site Beckerich

1. DATA SUMMARY	
Task & Data Manager	WP3 – T3.2
Availability	Private or Consortium

	A subset will be considered to be delivered as Open Data to enable and facilitate further research, upon validation from the Conseil Communal
Purpose of the data	The data collected by the in T3.2 is related to the execution of demo site activities. Data provided by REGEN tools and methods, in other tasks, will be centralized. Specific data might be generated for the execution of interventions, including data prepared for participation workshops to be organized with businesses and citizens.
Origin of the data	The dataset will comprise open data publicly available (including geoportail.lu), data handled by the municipality or other public bodies, data provided by private business (tbc) as well as data generated in other tasks of REGEN project.
Storage	The data will be stored either in the IT infrastructure of the municipality and in LIST IT infrastructure. In case an open data is delivered, it will be considered to use LIST's data management infrastructure and tools to deliver it, upon agreement from the Conseil Communal.
Data nature and size	The data considered in this dataset is geospatial data, and any other CAD drawings, as well as survey and questionnaires used to capture the feedback from the citizens.
Data utility	The data will be useful to the REGEN partners.
Dataset Specific Metadata	tbc
2. FAIR DATA	
Making data findable	tbc
Making data openly accessible	tbc

Making data interoperable	tbc
Increase data reuse	tbc
3. ALLOCATION OF RESOURCES	
Costs	tbc
Long term preservation	tbc
4. DATA SECURITY	
Data security, recovery and storage	Security measures will follow the policies of the municipality of Beckerich and/or LIST.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	tbc
Informed consent	tbc
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	tbc

T3.3 Laredo Demosite	
1. DATA SUMMARY	
Task & Data Manager	WP3, Task 3.3 Data manager: By the moment Tecnalía (Elena Usobiaga), in the future the NRO
Availability	Private
Purpose of the data	All data gathered from Laredo demo site concerning all demo activities (AFUR KPIs, Digital twin, surveys, etc.)
Origin of the data	Both reused and produced in the framework of the project (e.g. surveys)
Storage	Part of the data will be stored in the sharepoint, but the storage is not fully defined yet.
Data nature and size	Examples of information types: • Word, PDFs, XLS, CVS files

	• Geospatial information • Databases (e.g. Postgre/PostGIS) Expected size not defined yet.
Data utility	To all the partners participating in the demo.
Dataset Specific Metadata	Regeneration actions; accessibility; monitoring; digital twin; neighborhood regeneration office; public participation.
2. FAIR DATA	
Making data findable	documents (.pdf), tables (.xls, .csv), optionally semantic graph data (.xml)
Making data openly accessible	Will all data be made openly available? NO, only the original outputs of REGEN. source data will be referenced where possible. References to alternative sources of data will be provided.
Making data interoperable	Not defined yet.
Increase data reuse	Not defined yet.
3. ALLOCATION OF RESOURCES	
Costs	Not defined yet.
Long term preservation	Not defined yet.
4. DATA SECURITY	
Data security, recovery and storage	Not defined yet.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Not defined yet.
Informed consent	Yes
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	Not defined yet.

T3.4 Demo site Milan

1. DATA SUMMARY

Task & Data Manager	WP3 - T3.4 Demosite data collection leader: AMAT – Andrea Naldini andrea.naldini@amat-mi.it
Availability	Consortium (via precaution)
Purpose of the data	Evaluating, Monitoring, Benchmarking
Origin of the data	Mainly processed
Storage	REGEN website for public deliverables, Microsoft Teams, Milan database
Data nature and size	PDF, shp, kmz, dwg, skp, xls, docx
Data utility	Consortium level and external
Dataset Specific Metadata	Milan, Giaggioli, REGEN
2. FAIR DATA	
Making data findable	The data will be made accessible through standardized and open access protocols where possible, while others will have confidential access.
Making data openly accessible	Where possible, all datasets will be made openly available to the public. Certain datasets, particularly those containing confidential information for instance from construction companies or sensitive survey data from inhabitants, will not be openly accessible or anonymised for project's purposes. These restrictions are in place to comply with data protection regulations (e.g., GDPR) and to protect the privacy of individuals whose data has been collected. For restricted datasets, access will be granted on a case-by-case basis under specific conditions, such as signing a data use agreement that ensures confidentiality and data security.

Making data interoperable	Efforts will be made to use established and widely-accepted ontologies and vocabularies to ensure interoperability of the data. In cases where the use of uncommon or project-specific ontologies or vocabularies is unavoidable, mappings to more commonly used ontologies will be provided. This will be done to ensure that the data remains interoperable and can be integrated with other datasets.
Increase data reuse	Comprehensive documentation will be provided to facilitate data validation and reuse. The provenance of the data will be thoroughly documented using appropriate standards.
3. ALLOCATION OF RESOURCES	
Costs	Costs will be covered primarily through REGEN funding, institutional support, and potential collaborations.
Long term preservation	Decisions on what data to preserve and for how long are typically made based on research impact and compliance with institutional.
4. DATA SECURITY	
Data security, recovery and storage	The AFUR framework documents are stored on the LIST Teams, any development files as part of the framework will be backed up on the internal LIST git repository, which is secured on an internal network following strict LIST IT security protocols.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Data, like identities must be protected through anonymization techniques, ensuring confidentiality and ethical standards in data handling.
Informed consent	The citizens must be informed about how their data will be used and preserved, maintaining ethical standards and regulatory compliance.
6. OTHER ISSUES	

Other national/funder/sectorial/departamental procedures	Following national data management procedures prescribed by the Italian Data Protection Authority (Garante per la protezione dei dati personali) and sector-specific guidelines, such as those from the National Institute of Statistics (ISTAT).
--	---

T3.5 Demo site Dublin	
1. DATA SUMMARY	
Task & Data Manager	WP3 –T3.5 Demo Site Dublin City Council Dublin City Council – darby.mullen@dublincity.ie
Availability	Consortium
Purpose of the data	Developing Baseline Model Diagnosis of Dublin REGEN Demo site Informing energy models and project interventions
Origin of the data	Open sourced National data sets Ireland's Open Data Portal (http://data.gov.ie) Outputs generated from digital models developed under REGEN project
Storage	REGEN Website National data
Data nature and size	• PDF • shp • kmz • dwg • dxf • skp • xls • csv • doc • docx • jpeg • mp4 Size of files not defined yet expected to be 10kb < file > 10mb
Data utility	Consortium
Dataset Specific Metadata	Demosite, Ringsend, REGEN

2. FAIR DATA	
Making data findable	Data will be made accessible through a free and standardized access protocol where possible.
Making data openly accessible	Not all data will be made openly available due to licenses and restricted access including GDPR
Making data interoperable	Every effort will be made to ensure standardized ontologies and vocabularies are used. In cases where it is unavoidable and uncommon or project specific ontologies or vocabularies are used, mappings to more commonly used ontologies will be provided.
Increase data reuse	Documentation will be provided to validate data analysis and facilitate data re-use (e.g. methodologies, codebooks, data cleaning, analyses, definitions, units of measurement, etc.)? The provenance of the data be thoroughly documented using the appropriate standards?
3. ALLOCATION OF RESOURCES	
Costs	The costs for making data FAIR have not been determined yet. Costs associated with processing and managing data will be assigned to REGEN funding.
Long term preservation	Resources and expectations for long term preservation of data have not been determined yet.
4. DATA SECURITY	
Data security, recovery and storage	Data security measures have not been determined yet.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	All data will be anonymized
Informed consent	No Personal Data will be used
6. OTHER ISSUES	

Other national/funder/sectorial/departmental procedures	No
---	----

T3.6 Evaluation of demo sites	
1. DATA SUMMARY	
Task & Data Manager	WP3 - Demo site activities Task 3.6 - Evaluation of REGEN toolbox based on demo sites experience Task 3.6 is managed by DEMO with contributions from TEC and demo site clusters including KAW and VAR.
Availability	Dissemination level is PU - Public
Purpose of the data	The purpose of data collection, processing, and generation in Task 3.6 is to evaluate the effectiveness of the REGEN toolbox by analyzing its deployment in real-world demo sites. This includes assessing the tools' impact, user satisfaction, task completion rates, and adoption rates. Insights and feedback from this evaluation will be gathered to guide future improvements and the potential replication of the REGEN toolbox.
Origin of the data	Produced within the framework of the REGEN project
Storage	The data will be stored in repositories such as the Zenodo or Opendoar for public deliverables, Microsoft Teams for internal data, and other specified repositories to be decided for the data management plan.
Data nature and size	Expected formats include JSON, XML, CSV, and documentation in DOCX or PDF The size of the data is yet to be determined, and will depend on the volume of interactions and integrations.
Data utility	The data will be useful for consortium members to assess the results refine and improve the REGEN toolbox and for researchers and practitioners involved in

	urban regeneration and digital tool development.
Dataset Specific Metadata	NA
2. FAIR DATA	
Making data findable	./
Making data openly accessible	Not all data will be available openly due to privacy and security concerns. Certain datasets will be restricted to the consortium, as specified in the Grant Agreement. Some data may remain restricted due to legal and contractual reasons. A subset of the data will be anonymized and made openly accessible.
Making data interoperable	NA
Increase data reuse	Data documentation will be provided, including methodology, codebooks, and data-cleaning procedures to facilitate reuse if relevant
3. ALLOCATION OF RESOURCES	
Costs	Costs for making data FAIR are covered by the Horizon Europe grant.
Long term preservation	The data management plan has yet to determine what data to preserve and for how long.
4. DATA SECURITY	
Data security, recovery and storage	Data is stored securely in certified repositories with provisions for data recovery and secure transfer. Data security measures will be in place, including secure storage, data recovery plans, and transfer protocols for sensitive data. After discussion with other partners, certified repositories will be used for long-term preservation and curation.

5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Ethical considerations will be addressed in accordance with the approved ethics review and relevant guidelines. To ensure privacy and confidentiality, anonymization techniques will be implemented.
Informed consent	Informed consent will be obtained for any personal data collected, ensuring compliance with GDPR and other relevant regulations.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	No

6. Data collection in WP4

T4.1 Impact evaluation dataset	
1. DATA SUMMARY	
Task & Data Manager	WP3 – Task 3.1 + WP4 – Task 4.1 Data managers: NBK & LIST
Availability	Open
Purpose of the data	Describe the sustainability performances of the interventions tested at demo sites
Origin of the data	Produced in the framework of the project (AFUR evaluation results)
Storage	REGEN website (public deliverable)
Data nature and size	Table format (e.g. xls, csv) included in reports (pdf format) Size between 100 Ko and 10 Mo
Data utility	Consortium level: Promote interventions benefits or adjust interventions to avoid potential trade-offs Externally: Understand the impacts of interventions for their potential replication
Dataset Specific Metadata	Impacts type (e.g. global environmental, local social...), evaluated intervention (name, location, size...), calculation methodology, data source, data generator
2. FAIR DATA	

Making data findable	documents (.pdf), tables (.xls, .csv), optionally semantic graph data (.xml)
Making data openly accessible	No, only the original outputs of REGEN. source data will be referenced where possible. References to alternative sources of data will be provided.
Making data interoperable	Not defined yet
Increase data reuse	Methodological explanations and data sources will be documented as far as possible, following in particular LCA standards
3. ALLOCATION OF RESOURCES	
Costs	Not defined yet
Long term preservation	Not defined yet
4. DATA SECURITY	
Data security, recovery and storage	The AFUR framework documents are stored on the LIST Teams, any development files as part of the framework will be backed up on the internal LIST git repository, which is secured on an internal network following strict LIST IT security protocols.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	No personal data will be collected or used for the generation of AFUR results.
Informed consent	N/A
6. OTHER ISSUES	
Other national/funder/sectorial /departmental procedures	No

T4.3 AFUR replication assessment via 10 virtual cases

1. DATA SUMMARY

Task & Data Manager	WP4 T4.3 Data manager: [Tannya Pico]
Availability	Consortium
Purpose of the data	Data will be served in order to conduct: (1) sustainability impacts assessment, (2) aggregation of results and (3) benchmarking of adequate REGEN interventions from the catalogue. This will help refine the AFUR model to ensure its further exploitation and replication
Origin of the data	Data will then be gathered from 10 urban regeneration projects (either through municipalities or project promoters or project developers) Data origin will be from outside the project Data will be reused data from desk research
Storage	Data will not be stored publicly, as it includes data from organisations outside of the REGEN consortium Data will be stored on REGEN Microsoft Teams (Sharepoint)
Data nature and size	Excel files (1 per virtual case = 10) with numerical (quantitative) / text (qualitative) data Supplementary data may include .pdf, image files
Data utility	At consortium level The data will help refine AFUR (WP4) and feed into replication activities with WP5
Dataset Specific Metadata	“[urban regeneration project name]”, “interventions”, “impacts
2. FAIR DATA	
Making data findable	The dataset will not be accessible through a free and standardized access protocol Data from each virtual case will be organized using the same logic.

Making data openly accessible	The data will not be made openly available as it includes data from organisations outside of the REGEN consortium. Dataset is accessible for consortium only
Making data interoperable	Data will be organized using standard ontologies / vocabularies. In case of specific deviation, clear reference and explanation will be provided The data will come with clear definitions of the variables that can be understood across disciplines.
Increase data reuse	Data sources will be clearly indicated Clear guidelines will detail the data acquisition process, its cleaning, and coding.
3. ALLOCATION OF RESOURCES	
Costs	NA
Long term preservation	NA
4. DATA SECURITY	
Data security, recovery and storage	Project level measures to secure data stored in REGEN Microsoft Teams (Sharepoint)
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	The intended use of the collected data will be explicitly laid out to organizations from which the data will be collected. The ethics committee is expected to review deliverables that have been identified to need ethical and research integrity assessment
Informed consent	NA This data will not include personal data
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	NA

T4.4 Application of AFUR on 30 additional interventions	
1. DATA SUMMARY	
Task & Data Manager	WP4 – Task 4.4 Data managers: NBK & LIST
Availability	Open
Purpose of the data	Describe the sustainability performances of additional interventions (not tested at demo sites)
Origin of the data	Produced in the framework of the project (AFUR evaluation results)
Storage	REGEN website (public deliverable)
Data nature and size	Table format (e.g. xls) included in reports (pdf format) Size between 100 Ko and 10 Mo
Data utility	Consortium and externally: Understand the impacts of interventions for their potential replication
Dataset Specific Metadata	Impacts type (e.g. global environmental, local social...), evaluated intervention (name, location, size...), calculation methodology, data source, data generator
2. FAIR DATA	
Making data findable	documents (.pdf), tables (.xls, .csv), optionally semantic graph data (.xml)
Making data openly accessible	No, only the original outputs of REGEN. source data will be referenced where possible. References to alternative sources of data will be provided.
Making data interoperable	Not defined yet
Increase data reuse	Methodological explanations and data sources will be documented as far as possible, following in particular LCA standards
3. ALLOCATION OF RESOURCES	
Costs	Not defined yet
Long term preservation	Not defined yet
4. DATA SECURITY	

Data security, recovery and storage	The AFUR framework documents are stored on the LIST Teams, any development files as part of the framework will be backed up on the internal LIST git repository, which is secured on an internal network following strict LIST IT security protocols.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	No personal data will be collected or used for the generation of AFUR results.
Informed consent	N/A
6. OTHER ISSUES	
Other national/funder/sectorial/depart mental procedures	No

T4.5 Advocacy material and roadmap	
1. DATA SUMMARY	
Task & Data Manager	Work Package: WP4 - AFUR demo sites evaluation and replication Task 4.5 - Advocacy to support REGEN outcomes uptake, Task 4.5 is managed by DEMO
Availability	Dissemination level is PU - Public
Purpose of the data	Data from Task 4.5 will be used to create resources that promote and guide the adoption of REGEN project results, including best practices, training materials, and a roadmap. This will contribute to the EU taxonomy by addressing thresholds and DNSH principles.
Origin of the data	The project will generate data through workshops, surveys, stakeholder interviews, best practice analysis, feedback, and demo site implementation results.

Storage	The data will be stored in repositories such as the Zenodo or OpenDoar for public deliverables, Microsoft Teams for internal data, and other specified repositories to be decided for the data management plan.
Data nature and size	Expected formats include JSON, XML, CSV, and documentation in DOCX or PDF. Related data includes policy documents, best practice guides, training materials, advocacy roadmaps, feedback forms, and stakeholder engagement records. The size of the data is yet to be determined, and will depend on the volume of interactions and integrations.
Data utility	The data will be useful for consortium members, policymakers, urban planners, and other stakeholders involved in urban regeneration and decarbonization efforts.
Dataset Specific Metadata	Urban Regeneration, Digital Tools, User Experience, Performance Metrics, REGEN Toolbox, Evaluation, Feedback.
2. FAIR DATA	
Making data findable	NA
Making data openly accessible	Not all data will be available openly due to privacy and security concerns. Certain datasets will be restricted to the consortium, as specified in the Grant Agreement. Some data may remain restricted due to legal and contractual reasons. A subset of the data will be anonymized and made openly accessible.
Making data interoperable	NA

Increase data reuse	If relevant and needed, data documentation will be provided, including methodology, codebooks, and data-cleaning procedures to facilitate reuse.
3. ALLOCATION OF RESOURCES	
Costs	Costs for making data FAIR are covered by the Horizon Europe grant.
Long term preservation	The data management plan has yet to determine what data to preserve and for how long.
4. DATA SECURITY	
Data security, recovery and storage	Data is stored securely in certified repositories with provisions for data recovery and secure transfer. Data security measures will be in place, including secure storage, data recovery plans, and transfer protocols for sensitive data. After discussion with other partners, certified repositories will be used for long-term preservation and curation.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Ethical considerations will be addressed in accordance with the approved ethics review and relevant guidelines. To ensure privacy and confidentiality, anonymization techniques will be implemented.
Informed consent	Informed consent will be obtained for any personal data collected, ensuring compliance with GDPR and other relevant regulations.
6. OTHER ISSUES	

Other national/funder/sectorial/departmental procedures	No
---	----

6. Conclusions

The present document represents the first version of the REGEN Data Management Plan established in June 2024 (Month 6). It sets a benchmark to identify the actions that need to be implemented by the REGEN partners in order to fulfil the European Commission's requirements in terms of data management and accessibility of the research data.

The DMP is a living document and will be updated and further developed during the project's lifetime.