Start date of project: 1ˢᵗ December 2022

Duration: 48 months

D7.3 – initial data management plan

Closing the gap between fork and farm for circular nutrient flows

Deliverable details	
Work Package Title	**Project Management**
Task Number	**7.5**
Deliverable Number	**D7.3**
Deliverable Title	**Initial data management plan**
Version Number	**1.0**
Responsible Organisation	**agrathaer GmbH**
Author(s)	**Anita Beblek**
Due Date	**Wednesday, May 31, 2023**
Delivered Date	**Wednesday, May 31, 2023**
Reviewed by	**Dr. Stefan Karlowsky, Dr. Henrike Perner, Sven Möller**
Dissemination level	**PU - Public**
Please cite as	**Beblek, A., D7.3 Initial data management plan P2GreeN, 2023**
Contact person EC	**Sofia Pachini**

Contributing partners	
1.	AGRATHAER GMBH – Germany (Coordinator)
2.	LEIBNIZ-INSTITUT FUR GEMUSE- UND ZIERPFLANZENBAU GROSSBEEREN/ERFURT EV – Germany
3.	VILLE DE PARIS – France
4.	UNIVERSITY COLLEGE DUBLIN, NATIONAL UNIVERSITY OF IRELAND, DUBLIN – Ireland
5.	SVERIGES LANTBRUKSUNIVERSITET - Sweden
6.	COPENHAGEN BUSINESS SCHOOL - Denmark
7.	LUONNONVARAKESKUS - Finland
8.	NATIONAL UNIVERSITY OF IRELAND MAYNOOTH - Ireland
9.	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS - Greece
10.	BIOAZUL, SL - Spain

11.	CIP CITIZENS IN POWER - Cyprus
12.	ECOVILLAGE HANNOVER EG - Germany
13.	GOLDEIMER GEMEINNUTZIGE GMBH - Germany
14.	INSTITUT D'ARQUITECTURA AVANCADA DE CATALUNYA - Spain
15.	ICLEI EUROPEAN SECRETARIAT GMBH (ICLEI EUROPASEKRETARIAT GMBH) - Germany
16.	FUNDACION CENTRO ANDALUZ DE INVESTIGACIONES DEL AGUA - Spain
17.	SUN GLOBAL CHEMICALS SERVICES SL - Spain
18.	IRIDRA SRL - Italy
19.	ECOLE NATIONALE DES PONTS ET CHAUSSEES - France
20.	KOZGAZDASAG- ES REGIONALIS TUDOMANYI KUTATOKOZPONT - Hungary
21.	HAFENCITY UNIVERSITAT HAMBURG - Germany
22.	SUSTCHEM TECHNIKI SYMVOULEFTIKI ANONYMI ETAIREIA - Greece
23.	TRANSITION APS - Denmark
24.	TRIODOS BANK NV - Netherlands
25.	MOVERIM CONSULTING SPRL - Belgium
26.	SOCIEDAD AGRARIA DE TRANSFORMACION TROPS N 2803 DE RESPONSABILIDAD LIMITADA - Spain
27.	AGRI SMART DATA SL - Spain
28.	TOUCH DOWN GOTLAND AB - Sweden
29.	SANITATION360 AB - Sweden
30.	GOTLANDS BRYGGERI AB - Sweden
31.	AGUAS Y SANEAMIENTOS DE LA AXARQUIA SA - Spain
32.	VunaNexus - Switzerland

Project Number: 101081883
Project Acronym: *P2GreeN*
Deliverable 7.3

Initial data management plan

Content

Index of tables

1. Executive Summary

This document presents the initial Data Management Plan (DMP) of the P2GreeN project, Deliverable (D) 7.3 lead by agrathaer (AGR) as the coordinator. The aim of the initial DMP is to establish a consortium data management framework which establishes the general principles of data collection, data handling and data storage. AGR and IGZ as coordinator and co-coordinator together represent P2GreeN's Data Management Team. The Team is responsible not only for the Data Management but as well for the design and implementation of P2GreeN's dedicated innovation platform that will form an essential part of making P2GreeNs data FAIR (Findable, Accessible, Interoperable and Re-usable).

This DMP is based on a bottom-up process between the Data Management Team and all consortium members. The process started in Month (M) 3 with a first data management workshop by the Data Management Team, followed by a first consortium wide DMP and innovation platform workshop in M4 and a dedicated data management survey (Annex 1) in M4 and 5 with a very high participation by the consortium.

The DMP is a living document and will be updated as the implementation of the project progresses. A mid-term DMP is due in M34 and the final DMP in M46.

2. Data Summary

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

The initial results of the data management survey with all consortium partners conducted during M4 and 5 of the project have revealed that the pilot regions in WP 1 will re-use existing data that has been generated in previous projects and during the development process of the four different technology systems that will be operationalised in each of the three pilot regions. The existing data differ in each pilot region due to their different technology systems. The data will be used as the foundation for maturing the existing technological and systems innovation in the three pilot regions in Sweden, Germany and Spain from their current TRL 5 and 6 to TRL 8.

What types and formats of data will the project generate or re-use?

P2GreeN is built on a circular multi-sectoral systems approach with a high level of transdisciplinarity, including consortium members from different scientific disciplines, tech-providers, innovation brokers, industry, NGOs and public administration, Therefore, the data produced and collected in the project will be diverse. The initial data management survey gives a first initial overview on types and formats of data. According to this survey the following types of data will be generated: text, numeric data, cartographic and georeferenced data, software codes, pictures, videos and sound clips. Text data will be generated as plain text files, Microsoft Word Document, PowerPoint, Portable Document Format or Comma-separated values (TXT, DOCX, PPTX, PDF, CSV). Numerical data will be collected in Microsoft Excel (XLSX), CSV and SPSS (statistical package for social sciences) formats. The format of cartographic and georeferenced data is not defined yet at this early stage of the project, but common formats will likely be used as the minimum standard. These include ESRI's ArcGISMap Template (MXT) and documents (MXD), shapefiles (SHP) as well as GeoTIFF image (MXT, DIV, TIF). Further images, pictures, sound and video files will be generated in the format of Portable Network Graphic (PNG), Power Point (PPTX), digital video formats (MP4, MOV, WMV, AVI, WEBM, HTML5) and JPEG images (PDF, JPEG, TIFF).

What is the purpose of the data generation or re-use and its relation to the objectives of the

Project Number: 101081883
Project Acronym: *P2GreeN*
Deliverable 7.3

project?

The overall objective of P2GreeN is to foster a paradigm shift, from a linearly organised resource and nutrient system within the agri-food supply chain towards a circular material flow between urban and rural areas following the 3R principle "Reduce, Reuse, Recover". To safeguard achieving P2GreeN's objectives, the collection, curation, and preservation of data are of key importance. Hence there are three pillars related to data management, which will support meeting the overall objective:

1. A dedicated work package for scientific data acquisition in the whole P2GreeN project to generate and manage scientifically relevant and replicable data, WP 2;

2. The implementation of a network hub and repository for engagement, outreach and impact, WP 5, Task 5.1;

3. The establishment of the P2GreeN innovation platform to collect, host and serve the data, set of methods and tools that will be generate during the project and that will stay in operation for 5 years after the project, WP 7, Task 7.5.

The DMP will serve as a framework and supporting document for all the three pillars. Furthermore, to preserve the legacy of P2GreeN's objective and outcomes, a strong interconnection of the DMP with the innovation platform will be implemented by the Data Management Team consisting of the coordinator AGR and the co-coordinator IGZ.

To achieve the purpose of sound scientific data acquisition for the later filling of the P2GreeN innovation platform, the Data Management Team together with the co-leads of WP 2 have established the following protocol: i) for all data collected in WP 2, data collection templates will be designed, and ii) the use of these templates will be mandatory. As the innovation platform will be established in M12, all data collected in the meantime will be stored on the cloud-based software platform of the coordinator AGR.

What is the expected size of the data that you intend to generate or re-use?

At the current early stage of the project, a reliable statement on the scope of data cannot be made. The data storage capacity at AGR's cloud based collaborative platform (MS Teams and Sharepoint) is 25 Terabytes but during the design of the innovation platform an in-depth analysis of maximum data size per dataset will be conducted. Data size rules will be implemented and reported in the mid-term DMP.

What is the origin/provenance of the data, either generated or re-used?

In the first two years of the project, data will mainly originate from the three pilot regions: i) Gotland in Sweden, ii) the Hamburg-Hannover region in Germany, and iii) the La Axarquia region in Spain. In years 3 and 4, additional data will be generated within the 4 follower regions in Italy, Greece, France and Hungary.

Data from the three pilot regions will come from field and technology trials (urine dehydration in Sweden, multi-step faecal composting and Vuna process of urine nitrification in Germany, wastewater reclamation in Spain), local stakeholder engagement activities, and the analysis of local value chains and governance models.

The origin of data for the follower regions will be defined mainly through Task 5.2 and the task leader IRIDRA. Preparation work will start in M6 and data origin will be specified later in the mid-term DMP in M34.

To whom might your data be useful ('data utility'), outside your project?

The target groups to whom data shall be useful are manifold and already defined at a detailed level. In total, 9 target groups for whom the data shall be useful have been identified and are described in Table 1.

Table 1 *Identified target groups for data generated in P2GreeN.*

Target group	Description of the target group
Public Authorities	Managing authorities for circular bio-based nutrients, agricultural regulation, waste policies and environment, urban planning departments – at national, international (EU), and local level in the following countries: Sweden, Germany, Spain, Greece, Italy, France, Hungary
Professional networks	Urban and industrial networks, mainly active in the field of fertilisers (e.g. fertiliser organisations, fertiliser distribution companies and fertiliser use consulting agencies), but also innovator companies
NGOs	Organisations active in the field of promoting sustainable circular solutions that have a particular interest in sustainable farming and sustainable management of agricultural nutrients (e.g. Zero Waste Europe, WWF European Policy Office, European Sustainable Phosphorus Platform (ESPP), European Green Cities)
Policy makers at local, national and international level	Policy makers at EU Level (EU Commission)
	Policy makers at national and local level in the pilot regions countries (Sweden, Germany, Spain) and in follower regions (Greece, Italy, France, Hungary)
Farmers and farmers' associations	European farmers' associations (e.g. COPACOGECA) as well as regional and local farmers' associations from pilot region countries and follower regions (Sweden, Germany, Spain, Greece, Italy, France, Hungary)
Citizens in local communities	Rural and urban citizens in the pilot region countries (Sweden, Germany, Spain) and in follower region countries (Greece, Italy, France, Hungary)
Scientific community	Scientists and researchers, especially in the field of environmental protection, sustainable farming and sustainable management of agricultural nutrients
Wastewater treatment organisations	Wastewater treatment organisations and agencies in the pilot region countries (Sweden, Germany, Spain) and in follower region countries (Greece, Italy, France, Hungary)
Standardisation bodies	Standardisation bodies at EU level (e.g. CEN and CENELEC), as well as in the pilot region countries (Sweden, Germany, Spain) and in follower region countries (Greece, Italy, France, Hungary)

Project Number: 101081883
Project Acronym: *P2GreeN*
                                                                     Deliverable 7.3

3. FAIR data

3.1   *Making data findable, including provisions for metadata*

Will data be identified by a persistent identifier?

The results of the data management survey show that only a minority of the consortium is familiar with metadata standards, which is also a reflection on the composition of the consortium as "just" 1/3 of the consortium are research organisations. However, all research organisations do use the Digital Object Identifier (DOI) system for data and scientific publications. WP 2 together with the Data Management Team will define common metadata terms between M6 and M12 and specify how these will be implemented across the different types of datasets collected in the project, as this will be a prerequisite for programming the P2GreeN innovation platform. At this point in time, templates for metadata files are considered as being a valid option to ensure that all collected and generated data is storable and findable on the P2GreeN innovation platform. Identification of all metadata by means of a DOI will be evaluated amongst the consortium members once the innovation platform is set up and operating as it will become clearer then, if a standard DOI rule for all data is feasible due to the heterogeneity of P2GreeN's data. The outcome will form part of the mid-term DMP.

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

The decisions on common mandatory rules on metadata and use of metadata standards for all consortium members will be taken in the process of designing and programming the innovation platform in M12. However, the DMP endorses that metadata is fundamental for creating, preserving, handling and making use of the project data. To this end, WP 2 has already started to define common guidelines for data collection within the pilot regions, ensuring that data suppliers in the pilot regions submit a dataset together with the corresponding metadata file.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Standardised search keywords will be provided as a requisite for the functionality of the innovation platform as the platform will not only operate as a data platform for internal use within the consortium, but also serves as a linking hub for the P2GreeN network and host all tools developed by the project for external and widespread use. The Data Management Team will provide a suggestion by M12 and feedback will be collected in a participatory framework with all consortium members.

Will metadata be offered in such a way that it can be harvested and indexed?

Metadata will be offered via the innovation platform to allow for harvesting and indexing. The details will be developed in parallel with the selection of the technology stack of the innovation platform and the definition on the structure of the datasets that will form part of the internal (consortium members only) area of the innovation platform.

3.2   *Making data accessible*

Repository:

Will the data be deposited in a trusted repository?

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

(All 3 questions answered together) P2GreeN's main repository during the project lifetime and for 5 years after the project ends will be the P2GreeN innovation platform. All P2GreeN data and results will be hosted and made available through the innovation platform, hence it will have the same functions as other trusted data repositories. The P2GreeN innovation platform will be hosted by the co-coordinator IGZ as a scientific organisation. IGZ has the required expertise to ensure that scientific standards related to assigning a Digital Object Identifier are met and maintained. IGZ will ensure open access to all datasets including searchable digital identifiers and metadata. The innovation platform will be maintained and kept updated also for 5 years after the project ends to ensure that access to metadata and data by the public will be available as a project legacy. Budget for this is allocated through the Grant Agreement.

Data:

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.

The following outcomes including all relevant data which are part of the outcomes will be made openly available:

- 3 blueprints developed and validated for regionally balanced circular N & P flows in Gotland, Sweden, in the Hamburg-Hannover region, Germany, and in the La Axarquia region, Spain;

- 3 demonstrated and validated novel governance solutions at the interface of rural/coastal-urban/industrial environments in Germany, Sweden and Spain;

- 4 implemented governance solutions in the follower regions in Greece, Italy, France and Hungary;

- 3 tailor-made circular business models for the pilot regions;

- 4 feasibility studies for the follower regions;

- Legal scoping review and recommendations for legislative framework at national and EU level;

- Guidelines and Replication Action Plan for EU wide application of P2GreeN governance framework;

- 3 Life Cycle Assessments from the pilot regions with environmental, social and economic indicators.

The decision on how and to what extent all specific data, in relation to the 4 bio-based fertilisers tested and validated within P2GreeN, will be made openly available will form part of the IPR strategy. For this, a consensus is to be reached with the consortium members that provide the

innovative fertilisers, as these IPRs are their main business assets. It is therefore a legitimate interest by these consortium members to protect their IPRs.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

The P2GreeN DMP endorses the principle of making data openly available as soon as possible. However, as research organisations and technology providers are part of the consortium, situations may arise where embargos have to apply due to publishing constraints (e.g. dissertation thesis) or to seek IP protection, especially for the filing of patents. As part of the IPR strategy, due as Deliverable D7.1 in M15, a data handling agreement for the aforementioned situations will be developed, agreed on with all consortium members and implemented. This data handling agreement will establish the rules and rights and define common a standard procedure if a consortium member is seeking to apply for an embargo.

Will the data be accessible through a free and standardized access protocol?

The use of standardised communication protocols, in combination with persistent identifiers available on the innovation platform for all aforementioned P2GreeN results and their corresponding data, will ensure access to the datasets. The datasets will be automatically processable. For programming the innovation platform, the Data Management Team will safeguard the use of open source standard protocols to grant maximum access, specifying the conditions under which the data are accessible whilst maximizing the potential for re-use of the data.

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

The innovation platform will consist of two sections: an internal section for consortium members only and a public section that is available and accessible for the general public as well as for researchers and all P2GreeN stakeholders.

All consortium members will have unrestricted access to the internal section as well as the public section during the project and for 5 years after the project.

As the IPR strategy will be developed concurrently with the innovation platform, rules will be agreed amongst all consortium members which part of restricted-access data from a consortium member will be made available on the public section of the innovation platform, (rsp. on an individual case-by-case approval for specific requests to restricted-access data). The implementation and monitoring of these consented rules will be the responsibility of the Data Management Team.

How will the identity of the person accessing the data be ascertained?

Any data uploaded on the innovation platform will be linked to a unique registered user identity. The innovation platform will grant access to the internal section only to persons with a dedicated registered user identity. This is already a standard procedure for all data stored on the coordinator's software platform.

The identity of users that access data in the public section of the innovation platform will not be ascertained.

Is there a need for a data access committee (e.g. to evaluate/approve access requests to

personal/sensitive data)?

Rules of data access will be developed between the Data Management Team and WP 2 during the programming of the innovation platform and consented among all consortium members. The implementation and monitoring will be the responsibility of the Data Management Team.

Metadata:

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Metadata will be made openly available and free of restrictions under CC0 1.0. It will be recommended that users should give attribution to the metadata sources to increase outreach of the innovation platform and P2GreeN project results.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Data will remain available to the public on the innovation platform for a minimum of 5 years after the end of the project. Metadata and actual datasets will be stored in separate files. In this way, the metadata will remain findable and available even after the 5 years. Part of the final DMP, due in M46, will be the consortium's decision on the potential to safeguard open access to all public data, even beyond the 5 years, and how a permanent data accessibility as a P2GreeN legacy could be facilitated.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

As the process of the requirements description for the innovation platform is not yet finished, the architectural structure and technology stack are not yet defined as well. Hence, the question if a software will be needed to access and/or read the data is not sufficiently reliable to answer in this initial DMP. However, once the architectural structure and the technology stack have been defined and agreed upon, and if documentation or reference will be needed to access the data on the innovation platform, this will be established and made publicly available. In conjunction with this, it will then also be decided if the relevant software for access can be provided as open source.

3.3   *Making data interoperable*

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

As described previously, WP 2 and the Data Management Team will define data standards and common metadata vocabularies and will provide them as guidelines for all data production and collection. This will be done during M6-12 whilst the innovation platform is designed. P2GreeN unites many different disciplines, therefore, the guidelines will consider that bibliographic and administrative metadata can be standardised across disciplines whilst metadata on the process and content of research results often have a very discipline-specific structure and content. Still the data stored on the innovation platform will use standard vocabularies to enable inter-disciplinary access and exchange. The standards to be set for data collection, storage and accessibility will make it possible to comply with FAIR data criteria.

Project Number: 101081883
Project Acronym: *P2GreeN*
Deliverable 7.3

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? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

At this point of the project, it is not foreseen to use uncommon ontologies or vocabularies. If the requirement to use project specific ontologies or vocabularies arises, this will be discussed within the consortium as well as potential remedies. Only if consented amongst all consortium members and no other remedies can be found, project specific ontologies or vocabularies will be used. In this case, the Data Management Team will provide mappings and will openly publish the generated ontologies and/or vocabularies.

Will your data include qualified references1 to other data (e.g. other data from your project, or datasets from previous research)?

Qualified references will be included explaining their intent. Correspondingly, meaningful links between datasets that build on other datasets be it from previous research or other data from the P2GreeN project will be provided.

3.4 Increase data re-use

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.)?

For data collected in WP 2, the creation of readme files will be mandatory where standards-based metadata do not exist or where it is not feasible (due to the heterogeneity of data). The readme files will aim to provide specific information about a data file to help ensure that the data can be correctly interpreted by P2GreeN consortium members internally, because much of the collected data in WP 2 will feed into data processing in WPs 3, 4 and 5. Ultimately, this will also allow for data re-use by all P2GreeN stakeholders as well as interested scientists once the data will be made publicly available on the innovation platform. For the creation of readme files, the Data Management Team will provide guidelines following the principles of:

- ➢ Creation of readme files for logical clusters of data;

- ➢ Naming the readme file so that it is easily associated with the data file(s) it describes;

- ➢ Readme document as a plain text file, avoiding proprietary formats;

- ➢ Use of standardised date formats.

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Data that has been classified as "for public access" will be made freely available via P2GreeN's innovation platform. The classification process will be defined concurrently with the development of the innovation platform and the IPR strategy, as it constitutes and essential part of them.

---

[1] *A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)*

Will the data produced in the project be useable by third parties, in particular after the end of the project?

One of the main aims of the innovation platform is to make data and project outcomes (which will be multiple datasets) available to the public for re-use to foster dissemination and upscaling of the innovative systems solutions. Hence all data made available via the innovation platform will be useable by third parties for a minimum of 5 years after the project's end.

Will the provenance of the data be thoroughly documented using the appropriate standards?

Data collection and processing in WP 2 - as the main scientific data collection and processing WP - will be documented following W3C Recommendations, specifically the W3C PROV Data Model. The co-leads of WP 2 together with the Data Management Team will be responsible to ensure the compliance with W3C Recommendations for all data collected and processed in WP 2.

Describe all relevant data quality assurance processes.

The Data Management Team with support from the WP 2 co-leads will be responsible for assuring the quality of the data by making sure that datasets follow the FAIR principles included in this plan, and that data is updated on a regular basis.

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

4. Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

In Table 1 outputs generated by the project are listed. All outcomes will be made publicly available on the innovation platform in digital formats. A document download function will enable interested stakeholders, scientists and the general public to not only access the information, but also to save the files on individual data storage locations for independent use without any restrictions.

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

All project outputs will be governed under the P2GreeN FAIR framework developed together by the Data Management Team and the WP 2 co-leads. It is crucial for the upscaling ambition of P2GreeN that not only data but all project outputs available on the innovation platform adhere to the FAIR principles. This will safeguard the widespread uptake of P2GreeN's innovative and circular urban-rural nutrient flows as well as the associated novel governance solutions (blueprints).

5. Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g.

Project Number: 101081883
Project Acronym: *P2GreeN*
Deliverable 7.3

direct and indirect costs related to storage, archiving, re-use, security, etc.) ?

AGR has provided dedicated cloud computing resources via the MS Teams platform for interim data collection and processing until the innovation platform is in full function. It has a storage capacity of 25 Terabytes. IGZ will be responsible for setting up the innovation platform that will have sufficient data processing and storage capacity for the data collected and generated in P2GreeN.

In addition, WP 2 has allocated resources for data collection, data storage and implementation of the FAIR principle for all research outputs generated in WP 2 as part of each task.

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

All operative costs for data production, collection, processing and storage as well as for intra-consortium management are part of the allocated budget from the project to each consortium member. In addition, a dedicated budget for the innovation platform is allocated to IGZ in WP 7.

Who will be responsible for data management in your project?

The Data Management Team will be overall responsible for the data management. WP 2 co- and task leads will be responsible for all research data collected and produced and will liaise in frequent intervals with the Data Management Team. All WP co-leads will be responsible for safeguarding the implementation of and compliance with general data management principles created by the Data Management Team.

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

This point will be discussed and decided at a later stage of the project.

6. Data security

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

The collaborative MS Teams platform was set up by AGR and aims at ensuring a convenient, reliable and flexible framework for project collaboration, communication, monitoring and reporting until the innovation platform is operating in full function. It is available for all consortium members. Access is secured by individual user names and passwords, provided on demand by the project coordinator.

This platform is a secure tool for the storage and sharing of documents and all project information.

- ➢ All partners have the same level of information

- ➢ All meeting dates (general meetings, WP meetings) are available in a calendar

- ➢ The partners are requested to keep their own data up to date. This includes regular contact data such as identity, function, phone, email and postal addresses, and membership in Consortium Bodies and working groups, in particular WP teams

> ➢ A unique and centralised version of all project related documents is available to all partners. This presents multiple advantages such as:

- Full security of data is provided

- Multiple versions are avoided to prevent confusion

- Co-edition of MS Office documents is possible for maximal convenience and time efficiency

- Consortium members are requested to use the MS Word track mode if needed, in particular in the deliverables review process

- The versioning function allows tracing back of document elaboration and retrieval of possible errors

The innovation platform will adhere to the same security principles.

Will the data be safely stored in trusted repositories for long term preservation and curation?

Long-term storage of data in a trusted repository after the innovation platform's guaranteed continuous operation until 5 years after the project's end will be discussed and decided in year 4 of the project.

7. Ethics

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

In P2GreeN, there are no ethics and legal issues that have an impact on data sharing. At this point in time, it is also foreseeable that no such problems should arise.

Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

The management of data in the project will be in accordance with Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data and on the free movement of such data (GDPR). Therefore, all questionnaires will include information on the rights related to personal data and how P2GreeN will comply with all requirements of the GDPR.

8. Annex

DMP Survey

1. What do you consider an appropriate allocation of data management responsibilities within P2GreeN
   □ WP level   □ Task Level
   WP level   = data management rules apply collectively at WP level
   Task level = each task is responsible for data management

2. Are you familiar with metadata standards because you are using them in your work environment?
   □ Yes   □ No

3. If you are familiar with metadata standards, please name the metadata standards you are using
   _____

4. Should the DMP define common keywords for the data repository to help indexing, tagging and finding of data?
   □ Yes   □ No

5. If you voted yes in Question number 4, please indicate if you would be able to contribute to set up such a common glossary. Please state your name and the WP your are working in below
   _____

6. Do you consider common P2GreeN templates for data upload necessary?
   □ Yes   □ No

7. P2GreeN will produce intermediate and final results. Looking at your tasks and responsibilities do you consider it necessary that we determine a life cycle management plan?
   □ Yes   □ No
   life cycle management plan is a plan that would set out binding rules on how and when intermediate results and data would have to be archived and removed from our data/knowledge repository

8. If you answered Question number 7 with yes, please specify how often you will produce intermediate results and in which WP.
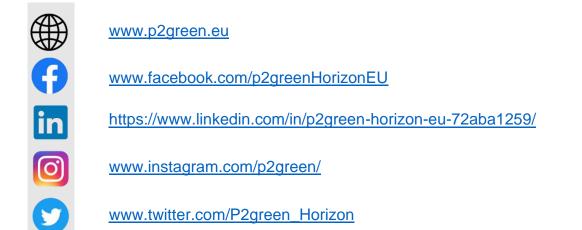   _____

Contact

Coordinator:

agrathaer GmbH

Eberswalder Straße 84

15374 Müncheberg

Germany

Tel.: +49 33432 82 141

info@agrathaer.de

www.p2green.eu

www.facebook.com/p2greenHorizonEU

https://www.linkedin.com/in/p2green-horizon-eu-72aba1259/

www.instagram.com/p2green/

www.twitter.com/P2green_Horizon