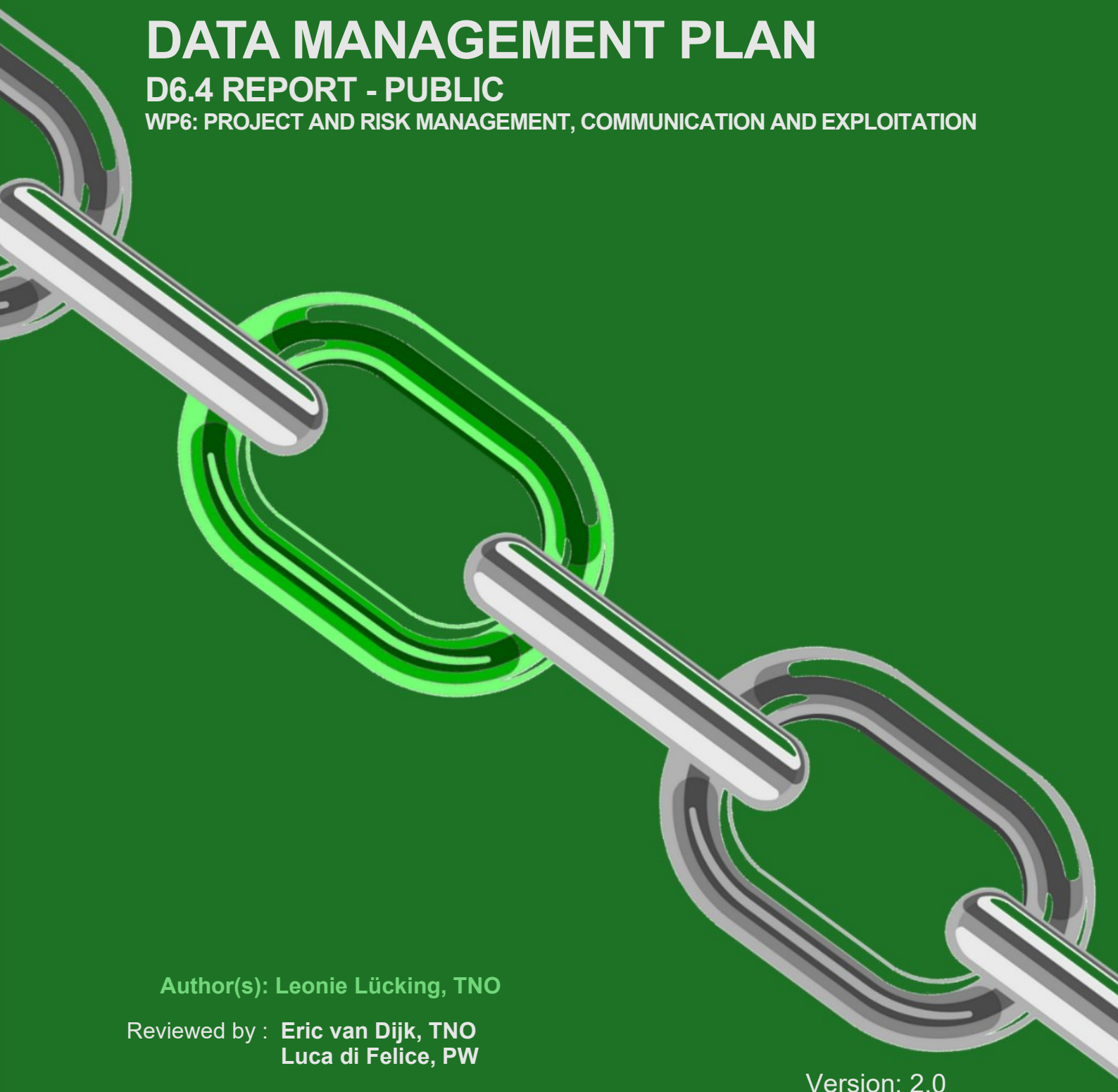
GreenSmith

PUBLIC

DATA MANAGEMENT PLAN

D6.4 REPORT - PUBLIC

WP6: PROJECT AND RISK MANAGEMENT, COMMUNICATION AND EXPLOITATION



Author(s): Leonie Lücking, TNO

Reviewed by : **Eric van Dijk, TNO**
Luca di Felice, PW

Version: 2.0

This Deliverable report D6.4 is part of the GreenSmith project, that was funded by CETPartnership, the Clean Energy Transition Partnership under the 2022 CETPartnership joint call for research proposal, co-funded by the European Commission (GA Nº 101069750) and with the funding organisations RVO (Netherlands), SWEA (Sweden) and MIMIT (Italy)

Version Management

VERSION	DETAILS
V1.0	Draft report for internal review (TNO)
V1.1	Draft review for review by consortium partners
V2.0	Final version

Disclaimer

Executive Summary

GreenSmith proposes a novel integration for the decarbonisation of the future low-carbon steel production. The project will de-risk and evaluate the application of the SEWGS technology to steel production routes integrating higher carbon processes (blast furnaces) to low carbon processes (hydrogen direct reduction) paving the way for implementation with the delivery of a basic engineering package for a TRL 8 GreenSmith plant.

The project will generate large amount of various data, e.g. experimental, modelling and engineering design, which will need to be handled according to the FAIR principle.

This Deliverable 6.4 report describes the strategy of the consortium for the handling, storage and distribution of generated data during the execution of the project. It also provides a description of the expected data generated by the consortium members in each WP. The proposed method for data management is based on the FAIR principle for findable, accessicble, interoperable and re-usability of data. Additionally, the data security and ethics are adressed.

Contents

List of abbreviations

DMP	Data Management Plan
EB	Executive Board
WP	Work Package

1. Introduction

The GreenSmith project will demonstrate operating reliability and technology-based innovations in demo-scale setting at TRL5 by producing hydrogen-rich gases. Being a time-ambitious and market close project with technical and operational challenges, GreenSmith utilises a Data Management Plan to ensure the longevity of the project outcomes and results.

This deliverable describes the generated data per work package (WP) and the principles and procedures of the Data management Plan (DMP): the FAIR method, the data security and ethics.

2. Data Summary

This section describes **the purpose** of the data collection and generation and the relation to the objectives of the project are highlighted. It also describes the **origin**, **type**, **format** and **size** of the data that is generated by the project activities, and explains how the **data will be used** in the project and who could be the **intended user** outside the project.

TABLE 1. List of the deliverables and their publication category

Deliverable	TItle	Publication category
D1.1	Operational report on Integrated BF-BOF and CH4 DRP Campaign	Confidential
D1.2	Operational report on H2 Campaign	Confidential
D1.3	Report on adsorbent post-test characterization	Confidential
D1.4	Exploitation of GreenSmith's experimental advancement	Public
D2.1	Selection of cost-effective optimal shapes	Confidential
D2.2	Demonstration of productivity increase of KC's Himago™ SEWGS adsorbent	Confidential
D3.1	Identified integration options	Confidential
D3.2	Optimisation of final integration schemes	Confidential
D4.1	TEA report for 2 selected cases	Confidential
D4.2	LCA report for 2 selected cases	Public
D4.3	Waste, by products and spent material management and recycling plan	Confidential
D4.4	Steelmaking Heat Map	Public
D5.1	Optimized reactor design report	Confidential
D5.2	Basic Engineering Package TRL8 plant	Confidential
D5.3	Report on characterization of gas BF and BOF gas streams	Confidential
D5.4	Report on preliminary Health, safety and environment analysis	Confidential
D6.1	Risk Management Plan	Public
D6.2	Communication and Dissemination Plan	Public
D6.3	Exploitation Plan	Public
D6.4	Data Management Plan	Public
D6.5	1st Annual progress report	Confidential
D6.6	2nd Annual progress report	Confidential
D6.7	Final Report	Confidential

The resulting data from the work within the GreenSmith project is reported in either confidential or public deliverable reports (see TABLE 1). The confidential reports are accessible for the consortium members, the CETPartnership and the three national funding agencies only, and are stored on a file SharePoint hosted by PW (see also section). The public deliverable reports will be made available through the project website. In order to disseminate the project results, the consortium strives to publish the project results via open access

peer-reviewed publications. Moreover, dissemination via interviews, non-scientific publications, courses and seminars will be performed. For all these dissemination actions that make use of confidential information, consent of the whole consortium has to be obtained using the procedure as set out in the CA.

The individual WP's are described below.

2.1. WP1 – Experimental Campaigns to Advance TRL

In WP1, the generation and acquisition of digital data are central to the main objectives of the TRL5 validation of the GreenSmith concept in the demo installation. A total of 2 campaigns will generate large amounts of raw data, forming the basis for the assessment of the performance of the columns and unit as a whole. The assessment involves detailed data analysis and modelling.

The raw data resulting from the demo operation will be of no direct use outside the project, as it relies heavily on background information as well as design details, which cannot be disclosed to the public and will only be described in confidential deliverable reports:

- D1.1 "Operational report on Integrated BF-BOF and CH_4 DRP Campaign"
- D1.2 "Operational report on H_2 DRP Campaign"
- D1.3 "Report on adsorbent post-test characterisation"
- D1.4 "Exploitation of GreenSmith's experimental advancement"

The data that is the basis for the deliverable reports will not be made publicly available as such, as the data reflects design and operational aspects that are considered IP sensitive and accordingly are classified as confidential. For dissemination purposes, the pilot operation as well as modelling results of the pilot operation will be described in general terms in the public space.

2.2. WP2 – Novel Functional Materials for Superior Performance

In WP2, data is generated to support the manufacturing of industrial functional materials and update the system model. Lab-scale testing generates data to consolidate the functional materials selection to the definition of the operating conditions. This information is input for the WP3 modelling. The lab-scale performance of the materials is evaluated under realistic operating conditions. The raw data resulting from these activities are confidential, as it relies heavily on background information from the partners, which cannot be disclosed to the public and will only be described in confidential deliverable reports:

- D2.1 "Selection of cost-effective optimal shapes"
- D2.2 "Demonstration of productivity increase of KC's Himago™ SEWGS adsorbent"

The data that is the basis for the deliverable reports will not be made publicly available as such, as the data reflects design and operational aspects that are considered IP sensitive and accordingly are classified as confidential. For dissemination purposes, information on industrial functional materials will be described in general terms in the public space.

2.3. WP3 – Conceptual Process Design to Optimise Integration Schemes

In WP3, data is generated to optimise of the GreenSmith concept for the chosen process integrations. The results will be the basis for the analysis in WP4 and the design in WP5 and provide input to WP1 and WP2.

The data will be described in confidential deliverable reports:

- D3.1 "Identified integration options"
- D3.2 "Optimisation of final integration schemes"

In order to disseminate the project results, it is considered that some of the information contained in these deliverables reports, process performance, etc. may be made public in open access peer-reviewed publications. The relevant data, equations and definitions developed to generate/process data in these publications, will be made public as supplementary information in such publications or in a public repository.

2.4. WP4 – Cross-Cutting Dimensions Analysis & Evaluation

In WP4, data is generated to support the evaluation of the GreenSmith concept in techno-economic terms, as well as life cycle characteristics, comparing the INTIATE concept to the next best available alternative technology and base-case scenarios.

The data will be described in confidential deliverable reports:

- D4.1 “TEA report for 2 selected cases”
- D4.3 “Waste, by products and spent material management and recycling plan”

In order to disseminate the project results, some of the information contained in these deliverables reports, costing, etc. will be made public in open access peer-reviewed publications. The relevant data, equations and definitions developed to generate/process data in these publications, will be made public as supplementary information in such publications or in a public repository.

Besides these confidential reports, a large part of the research data is made publicly available by means of public deliverables. These deliverables serve to inform the different stakeholders on the project results:

- D4.2 “LCA report for 2 selected cases”
- D4.4 “Steelmaking Heat Map”

These deliverables will be published on the project website. In addition, the assumptions and methodology for the process modelling and the techno-economic assessment will be made publicly available on a public repository so that the assessment can be retrieved by other researchers.

2.5. WP5 – Basic Engineering Design Package of a TRL8 Plant

In WP5, data is collected and generated to support the design of a TRL8 STEPWISE plant. This data include unit design, functional descriptions of components and machinery, stream compositions on-site and heat and mass balances. The majority of the generated data is intended for internal use in the design process. The format of the data varies with the intended use of the data. All design information, as well as knowledge on the operational characteristics of equipment and materials, is used in risk assessment analyses to be performed during WP5. These risk assessments include HAZID[1] assessments.

Data to be shared with 3rd parties concern:

- The technical and functional specifications of equipment and machinery that will be shared with suppliers to estimate the cost of equipment.

The data generated in the different design and engineering phases are confidential, as it relies heavily on background information from the partners as well as design details. The confidential information generated in the different design phases cannot be fully disclosed to the public and will be described in confidential deliverable reports:

- D5.1 “Optimized reactor design report”
- D5.2 “Basic Engineering Package TRL8 plant”
- D5.3 “Report on characterization of gas BF and BOF gas streams”
- D5.4 “Report on preliminary Health, safety and environment analysis”

The related data will not be made publicly available as several measurement and design aspects are considered IP sensitive and accordingly are classified as confidential. For dissemination purposes, the design and engineering will be described in general terms in the public domain.

[1] HAZID = HAZards IDentification

2.6. WP6 – Project Management, Communication and Exploitation

In WP6 data collection is aimed to ensure efficient and timely dissemination, communication & result exploitation. To do this, multiple activities are possible, e.g. guest lectures, massive online open courses (MOOCs), workshops and/or seminars. Personal data related to the registration of participants will be collected: first name, last name, email address, payment information and dietary information (if the event requires). Registration will proceed via the GreenSmith website, via an informed consent procedure. This data will not be made available for parties external to the consortium.

Further, no other data collection or generation will occur, other than processing of the data generated in the other WP's for the purpose of reporting to the CETPartnership and the funding agencies.

The communication and dissemination plans and the scientific learnings from the events will be shared through the following publicly available deliverable reports. These reports will be published on the GreenSmith website:

- D6.1 "Risk Management Plan"
- D6.2 "Communication and Dissemination Plan"
- D6.3 "Exploitation Plan"
- D6.4 "Data Management Plan"

A public summary of the annual reports will be posted to the website while the full reports listed below will remain confidential.

- D6.5 "1st Annual report"
- D6.6 "2nd Annual report"
- D6.7 "Final report"

3. Fair Data

Making provisions in the project to ensure that the research data generated will be **Findable**, **Accessible**, **Interoperable** and **Re-usable** (FAIR) is one of the main objectives of the DMP.

3.1. Making data findable, including provisions for metadata

This section describes approaches that will be followed by the partners in various work packages to ensure that the data produced and/or used is **discoverable**, **identifiable** and **locatable**.

The individual partners will make effort to ensure sensible naming of project related documents and that all documents are comprehensible. Reports and publications will contain information on the author(s), date and place of publication, as well as lists for version management, contents, abbreviations, references and acknowledgements.

For experimental campaigns, testing plans are recorded and raw data is collected and stored such that it can be retrieved and linked to the experimental planning. Following the data acquisition, data analysis files are accompanied by metadata that describe the type of experiment, the goal, main observations and conclusions. The results from the data analysis is commonly reported in the deliverable reports.

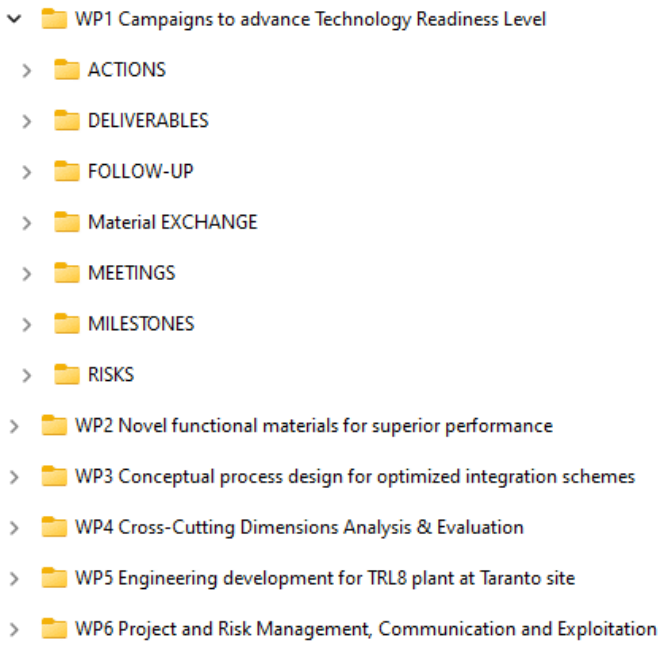


FIGURE 1 - Data storage structure

Simulation protocols also contain metadata, describing the goal of the simulations, what specific data is used, the conditions, input and output data files.

Each partner assures that their conventions for naming file and folders allows retrieving the data during and following the project. The data shared with the entire consortium is centrally stored on the server of the consortium partner Paul Wurth and accessible to all partners. The folder structure reflects the project structure in terms of WP number and category (see FIGURE 1). The naming of individual files reflects the relevant WP, Deliverable or Task number, responsible partner and date/version number (e.g. GreenSmith_WP6_D6.4_Data-Management-Plan_TNO_v1.0.pdf).

For publications in open access, DOI will be assigned (journal/ conference papers and supplementary materials), keywords will be listed in the reports/publications and acknowledgements will be made to the project.

3.2. Making data openly accessible

Data related to Public Deliverables will be openly available as by default. The data related to IPR protection or to relevant provisions made in the Consortium Agreement will be eligible to be s hared under the defined conditions.

According to the type of data and its level of confidentiality, the data will be made accessible on the project's communication channels (i.e. project website) or recognized repositories (for raw data of publication or data sets) which uses standard communications protocols. This will be defined during the project once the data is better defined and known. To ensure the safety of the data, the involved project partners will use their available local file servers to periodically create backups of the relevant materials. All other relevant documentation created during the project (i.e. deliverables) will be archived and preserved in PW SharePoint repository. It allows users to store, edit and share files within the project consortium.

All research data and associated material will be preserved for at least 5 years after the end of the project accoding to standard practices i.e European Commission standard practice for european projects.

The coordinator of the GreenSmith project will be responsible for the data management.

3.3. Making data interoperable

This section describes the provisions made to make the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc.

The data is made interoperable by ensuring that standard formats are used when saving. e.g. *.m for Matlab files, *.xlsx for Microsoft excel files, *.docx for MS Word files.

During writing and subsequent reviewing of project reports and publications, easy comprehension to the reader will be ensured. Deliverables and publications will be written in English. Standard format and vocabulary will be followed whenever possible to ensure the interoperability of the data. Where appropriate, list of abbreviations, nomenclature and key definitions will be provided, and tabulated data will be included in appendices. All reports and publications will have a summary.

3.4. Increase data re-use (through clarifying licences)

This section describes provisions made to ensure **efficient and wide re-use** of the project data, including quality assurance, data access licensing and any limitations on the length of time for which the data will remain re-usable.

The GreenSmith project results will preferably be published in open access unless the data is subject to confidentiality restrictions as described in the CA. Internal reports will be peer-reviewed by the relevant project partners and the Project Coordination Team. External publications will be reviewed by the publisher. Publications in the open access will be licensed via the research institution open repository and made available after publication for re-use. The data will be reusable until evidence of new data or information emerges.

The personal data collected in WP6 will not be re-used for anything other than the purposes of the project. This is due to GDPR and is established through informed consent. The data will be stored only as long as they are needed for the project purposes, and all personal data will be deleted at the end of the project.

4. Other research outputs

No other research output than listed in the previous sections are foreseen for the time being. In case other outputs are generated during the implementation of the project, the DMP will be updated accordingly.

5. Allocation of resources

The research data management in the project is overseen by TNO as project coordinator. Each partner is responsible for allocating resources to ensure efficient and secure storage of data generated in the project in accordance with institutional practices and regulations. Partners will use their institutional repositories free of charge. To cover the costs associated with the Article Processing Charges (APCs) for publications in open access journals, a provision was made in the partner's budget (see GA).

The GreenSmith project will use standard tools and a free of charge research data repository. The costs of data management activities are limited to project management costs and will be covered by allocated resources in the project budget of the coordinator and all partners..

6. Data security

This section describes the provisions put in place for data security, including data safe storage, data recovery, and transfer of sensitive data.

All GreenSmith partners will follow institutional policies and recommendations for data storage and backup. During the GreenSmith project, generated data intended for sharing between the partners will be stored in a secure database in Microsoft's Teams, hosted and managed by PW. The data are located within the European Union (Ireland and the Netherlands). Access to the information is possible worldwide through an invited Teams account, but always based on Multi-Factor Authentication. Microsoft offers standard facilities to ensure the confidentiality, integrity and availability of the information. 'In transit' and 'at rest' encryption is used for SharePoint Online. Additional information about the various certifications of the Microsoft Cloud services is available at Microsoft Trust Center[2]. Retention time for curated datasets is the same as other project materials at TNO, by default twenty years.

7. Ethics

In accordance with Article 14 (Ethics and Values) and Annex 5 of the Grant Agreement, the beneficiaries will carry out the project in line with the highest ethical standards and the applicable EU, international and national law on ethical principles.

In short, partners shall not perform activities that are prohibited in all EU Member States or prohibited in the Member State of the specific partner. Such activities are also not eligible for funding.

[2] Microsoft Trust Center Home, https://www.microsoft.com/en-us/trust-center

Partners must also pay particular attention to the principle of proportionality, the right to privacy, the right to the protection of personal data, the right to the physical and mental integrity of persons, the right to non discrimination, the need to ensure protection of the environment and high levels of human health protection.

Partners must ensure that the activities under the action have an exclusive focus on civil applications and do not involve human embryonal and/or cloning activities.

In addition, the beneficiaries must respect the fundamental principle of research integrity as set out in the European Code of Conduct for Research Integrity. This implies compliance with the following principles:

- **reliability** in ensuring the quality of research reflected in the design, the methodology, the analysis and the use of resources
- **honesty** in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair and unbiased way
- **respect** for col leagues, research participants, society, ecosystems, cultural heritage and the environment
- accountability for the research from idea to publication, for its management and organisation, for training, **supervision** and mentoring, and for its wider impacts and means that beneficiaries must ensure that persons carrying out research tasks follow the good research practices including ensuring, where possible, openness, reproducibility and traceability and refrain from the research integrity violations described in the Code.

For any arising questions or uncertaintises during the project the TNO data protection officer is aavailable to support the project

8. Conclusions

The GreenSmith DMP provides the framework for publishing the project research data in the open access repositories and making provisions to ensure the data is FAIR (i.e. Findable, Accessible, Interoperable and Re-usable). Whenever, during the execution of the project, there is a cause to update the DMP, a new version will be released.