CALiMERO

IMPROVING BIO-BASED INDUSTRIES LIFE CYCLE SUSTAINABILITY

D7.4. Intermediate version of the Data Management Plan

Due date of submission: 31/12/2023

Actual submission date: 03/01/2024

Funded by
the European Union

TABLE OF CONTENTS

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

PROJECT INFORMATION

Project full title: Industry CAse Studies AnaLysis To IMprove EnviROnmental Performance And Sustainability Of Bio-Based Industrial Processes

Acronym: CALIMERO

Call: HORIZON-CL6-2021-ZEROPOLLUTION-01

Topic: HORIZON-CL6-2021-ZEROPOLLUTION-01-06 - Increasing the environmental performance of industrial processes in bio-based sectors: construction, woodworking, textiles, pulp and paper and bio-chemicals

Start date: 1st July 2022

Duration: 48 months

List of participants:

Partner No.	PARTICIPANT ORGANIZATION | ACRONYM
1 (Coord.)	Contactica | CTA
2	WeLOOP | WELOOP
3	European Cellulose Insulation Association | ECIA
4	Swedish Environmental Research Institute | IVL
5	Neovili | NEOVILI
6	Cesefor | CESEFOR
7	Luxembourg Institute of Science and Technology | LIST
8	Technical University of Denmark | DTU
9	Techtera | TECHTERA
10	Essity | ESSITY
11	BIM Kemi AB | BIMKEMI
12	Ereks garment| EREKS

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

DELIVERABLE DETAILS

Document Number:	D7.4
Document Title:	Data management plan
Dissemination level	PU – Public, fully open,
Period:	PR1
WP:	WP7. Project management
Task:	T7.2. Data Management
Author:	CONTACTICA S.L.
Abstract:	This document corresponds to the intermediate version of Deliverable "Data Management Plan". It covers the description of how research data will be collected, processed, monitored, and catalogued during the CALIMERO project lifetime. For each dataset, it describes the type of data and their origin, the related metadata standards, the approach to data sharing and target groups, and the approach to data archiving and preservation, taking into account the need to balance openness and protection of scientific information, commercialisation, Intellectual Property Rights (IPR), privacy concerns and security. The CALIMERO Data Management Plan will be updated periodically.

Version	Date	Description
V0	03/01/2024	Initial version

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1 INTRODUCTION

This document corresponds to the first version of the Data Management Plan. It covers the description of how research data will be collected, processed, monitored, and catalogued during the CALIMERO project lifetime. For each dataset, it describes the type of data and their origin, the related metadata standards, the approach to data sharing and target groups, and the approach to data archiving and preservation, taking into account the need to balance openness and protection of scientific information, commercialisation, Intellectual Property Rights (IPR), privacy concerns and security. The information is organised by Work Packages (WP) and corresponds to the Data Management Plan aspects covered in some Guidelines on FAIR Data Management (in general terms, research data should be "FAIR", that is findable, accessible, interoperable, and re-usable). Information at this stage of the project has been gathered from Work Package Leaders (WPL). The CALIMERO Data Management Plan will be updated periodically.

2 DATA MANAGEMENT PLANS PER WORK PACKAGE

The data collection/generation within the project will be done with the purpose of achieving CALIMERO objectives. The thorough collection, evaluation and storage of experimental data will ensure reproducibility on the one hand, and traceability on the other hand, if experimental adjustments become necessary.

The non-confidential data sets, i.e., data that do not compromise the IPR from the partners, will be in open access through the Zenodo repository (https://zenodo.org/communities/calimero).

CALIMERO project will use this repository due to its characteristics:

- **Safe** — your research is stored safely for the future in CERN's Data Centre for as long as CERN exists.
- **Trusted** — built and operated by CERN and OpenAIRE to ensure that everyone can join in Open Science.
- **Citable** — every upload is assigned a Digital Object Identifier (DOI), to make them citable and trackable.
- **No waiting time** — Uploads are made available online as soon as you hit publish, and your DOI is registered within seconds.
- **Open or closed** — Share e.g., anonymized clinical trial data with only medical professionals via our restricted access mode.
- **Versioning** — Easily update your dataset with our versioning feature.
- **GitHub integration** — Easily preserve your GitHub repository in Zenodo.
- **Usage statistics** — All uploads display standards compliant usage statistics

Zenodo helps researchers receive credit by making the research results citable and through OpenAIRE integrates them into existing reporting lines to funding agencies like the European Commission. Citation information is also passed to DataCite and onto the scholarly aggregators.

1.1 WP1. CHALLENGES RELATED TO BIO-BASED INDUSTRIES

Work package	WP1. CHALLENGES RELATED TO BIO-BASED INDUSTRIES
1. Data summary	
1.1. Purpose of data collection/generation	Environmental, social and economic data is collected from industrial partners within CALIMERO. The aim is to cover a set of selected processes for each of the five bio-based sectors addressed within CALIMERO.

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1.2. Relation to project objectives	The collected data are crucial for almost all of the tasks of CALIMERO, because they are needed for every task relying on the Life Cycle Assessment (LCA) methodology and consequently tasks of WP1, WP2, WP3, WP4 and WP5 that rely on the results of LCA to emit recommendations for the bio-based sectors in Europe. Nevertheless, the collected data will directly allow reaching specific objectives: - identifying gaps in data, e.g., due to the lack of knowledge of industrials on environmental, social and/or economic aspects of studied processes (task 1.1) - identifying environmental, social and economic hotspots (task 1.2) - performing a holistic Life Cycle Sustainability Assessment (LCSA) of selected bio-based processes (task 2.1) - simulating most relevant industrial processes for the five bio-based sectors (task 2.2)
1.3. Types/format of data	Most of the collected data is confidential. They are stored securely by responsible partners in Excel spreadsheet format, following the Data collection protocol and Data collection sheets developed in task 1.1 and presented in deliverable 1.1. Once collected and compiled, non-confidential data will be identified and indexed in a FAIR format. Confidential data will be shared with the CALIMERO partners in a format allowing to protect the confidentiality while completing the LCA- related tasks.
1.4 Origin of data or reuse of existing data	Environmental and economic data originates from industrial partners of CALIMERO as well as a range of existing databases. Environmental data may be obtained from widespread databases (ecoinvent and GaBi). Social data is collected from an online survey realised using the surveymonkey platform (task 1.1, 1.2 and 2.1) as well as the Social Hotspots Database (SHDB) and the Product Social Impact Life Cycle Assessment (PSILCA) database. Data and data gaps could be completed and complemented with data obtained from literature reviews throughout CALIMERO.
1.5 Scale of data	While the number of single data points is difficult to estimate, collected data should cover approximately 15 to 20 processes across the five bio-based sectors.
1.6 Data utility	All collected data will be useful throughout the CALIMERO project, for partners involved in the tasks highlighted above, and particularly for partners involved in technical developments including environmental LCA, social LCA and Life Cycle Costing within WP1, WP2, WP3, WP4 and WP5. Outside of the CALIMERO project, non-confidential data will be useful for LCA practitioners, as well as industrials from the bio-economy, sustainability researchers and policymakers concerned by the bio-economy, especially in Europe.
2. FAIR Data	
2.1. Making data findable, including provision for metadata	
Facilitating findability	Data will be mainly disclosed through scientific reporting (articles, book chapters, etc.) and task deliverables, which any user can find with simple keywords-based search on Google scholar, Scopus and other scientific databases, including EU repositories that can prompt towards finding CALIMERO's information. Otherwise, since data will also be included in internal repositories (accessible e.g., through the project website), search options by keywords and filters might be included to facilitate and speed up the access to data.
Identifiability of data	Most data will remain confidential, but some publications can result after using the outcomes of WP4 and will do have a DOI.

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Versioning	Versioning will be used.
Metadata usage	No metadata format is expected to be used
2.2 Making data openly accessible	
Accessibility	To be decided
Method of availability	Public deliverable will be shared in the website of Calimero
Methods/software needed to access the data	Internet browser and internet connection
Access control	No restriction of access
2.3 Making data interoperable	
Interoperability	Data will be published to ensure they are findable, accessible, and interoperable and reusable internally (not by public), given the confidentiality.
2.4 Data reuse and quality	
Licensing	No
Data availability and embargo	No
Reuse restrictions	Not applicable
Data retention	At least 5 years after the end of the project
Data quality	To be decided
3 Other research outputs	
3.1. Plan for the management of other research outputs	To be determined
3.2. FAIR outputs	To be determined
3 Allocation of resources	
3.1 Costs	Costs are included in project budget
3.2 Data management responsibilities	Task leaders will be responsible of the data collected to perform the work of each specific task
3.3 Costs of preservation	Costs are included in project budget. Long-term storage of data costs will be assumed by each partner.
4 Data security	
4.1 Data security	Data will be stored in online servers with backup, recovery and secure storage services.
5 Ethical aspects	
5.1 Ethical aspects	NA
6 Other aspects	
6.1 Other aspects	NA

1.2 WP2. ANALYSIS OF CASE STUDIES FROM THE TARGET BIO-BASED SECTORS

Work package	**WP2. ANALYSIS OF CASE STUDIES FROM THE TARGET BIO-BASED SECTORS**
1. Data summary	
1.1. Purpose of data collection/generation	The purpose of data collection is to assess chosen case studies in the five bio-based sectors.
1.2. Relation to project objectives	The data collection in WP 2 is related to O1 “To identify main barriers and incentives to apply life cycle thinking sustainability approaches and source of impacts in the target bio-based sectors” and, O2 “To define reference case studies and identify levers to improve life cycle sustainability assessment methodologies and sustainability performances”
1.3. Types/format of data	The data collected from industrial partners is confidential. The data collected from databases and literature is not confidential.

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

	The data will be collected within spread sheets.
1.4 Origin of data or reuse of existing data	There are three origins for the collected data; (1) Data collected from industrial partners, (2) Data from databases like Ecoinvent, SHDB and PSILCA, (3) Data collected from literature. The latter two will be reused since they already exist.
1.5 Scale of data	The total amount of data can be estimated as 20 MB.
1.6 Data utility	The data will be useful for all the technical WPs of CALIMERO
2. FAIR Data	
2.1. Making data findable, including provision for metadata	
Facilitating findability	The data other than confidential data will be available in project deliverables and scientific open access articles.
Identifiability of data	The scientific articles will have DOIs. Do you expect to make use of identification mechanisms such as Digital Object Identifiers (DOIs)?
Versioning	It is expected to use versioning.
Metadata usage	No metadata usage is expected.
2.2 Making data openly accessible	
Accessibility	Confidentiality agreements are signed with industrial partners, therefore the data they share will be kept private. The results will be made openly available in a way that the confidentiality agreements are kept.
Method of availability	The openly available data will be made available in open access articles.
Methods/software needed to access the data	No specific software/tools will be needed to reach the openly available data.
Access control	There will be no restriction of access for openly available data.
2.3 Making data interoperable	
Interoperability	To be defined.
2.4 Data reuse and quality	
Licensing	Data will not be released under any specific licence. However, data incorporated in published articles will rely on the specific licence of use foreseen by the publication journal.
Data availability and embargo	An embargo will not be used and the data will be readily available.
Reuse restrictions	There won`t any restrictions for the reuse of the data.
Data retention	The confidential data will be retained at least 5 years after the closure of the project. There is no retention time for openly available data.
Data quality	To be decided.
3 Other research outputs	
3.1. Plan for the management of other research outputs	Not applicable
3.2. FAIR outputs	Not applicable
3 Allocation of resources	
3.1 Costs	To be estimated. In principle there will not be costs additional to those already foreseen within the CALIMERO budget
3.2 Data management responsibilities	Task leaders will be responsible of the data collected to perform the work of each specific task
3.3 Costs of preservation	No additional cost is expected.
4 Data security	
4.1 Data security	The confidential data are kept in shared folders which are only shared by IVL and the industrial partner. The rest of the data is kept in IVL`s data storage which is secured.
5 Ethical aspects	

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

5.1 Ethical aspects	Not applicable
6 Other aspects	
6.1 Other aspects	Not applicable

1.3 WP3. TOOLBOX/METHODOLOGICAL DEVELOPMENT FOR THE LCSA OF BIOBASED SECTORS

Work package	**WP3. TOOLBOX/METHODOLOGICAL DEVELOPMENT FOR THE LCSA OF BIOBASED SECTORS.**
1. Data summary	
1.1. Purpose of data collection/ generation	Process-based, literature and statistical data will be collected in various formats and elaborated to generate new data and information oriented to improve the life cycle sustainability assessment (LCSA) framework for biomass-based systems analysis.
1.2. Relation to project objectives	New data will be generated in order to estimate missing characterization factors for LCSA (objective of Task 3.1). Existing data will be collected (and new knowledge generated) to accomplish the objective of Task 3.2 (Integration of circular economy and criticality aspects into the LCSA framework). Statistical and literature data and information will be retrieved and used to improve the LCI calculation of GHG including temporal dimension (objective of Task 3.3). Qualitative knowledge on social issues and data from existing social aspects indicators will be used to integrate a socio-economic perspective in the LCSA framework (objective of Task 3.4).
1.3. Types/format of data	Different types of data will be collected and used as input for accomplishing the objectives of the WP3, among which the generation of new data in the form of characterisation factors. Those data will be mainly retrieved in spreadsheet format and may encompass both qualitative (nominal and/or ordinal) and quantitative (discrete and/or continuous) data, entered in physical and/or monetary units, as well as or even dimensionless and with georeferenced information. Conceptual or methodological knowledge from the literature will also be collected and elaborated to generate new data, but to a lesser extent than the direct retrieval of data from previous evidence and studies.
1.4 Origin of data or reuse of existing data	Data will be collected from the open access and proprietary literature (both scientific and grey), from commercial and free of access databases, it will be generated by the use of modelling tools as simulation outputs as well as from partners' surveys.
1.5 Scale of data	N/A, impossible to estimate at this stage.
1.6 Data utility	Data collected as such from the literature or other sources (with reference to 1.4) will be used by the project partners to develop models, perform analyses and quantify indicators in the four tasks of the WP3. Data generated by the activities in those tasks will be possibly all published or stored in open access databases for further exploitation by the broader scientific community and other interested public or private audiences.
2. FAIR Data	
2.1. Making data findable, including provision for metadata	
Facilitating findability	Data will be mainly disclosed through scientific reporting (articles, book chapters, etc.) and task deliverables, which any user can find with simple keywords-based search on Google scholar, Scopus and other scientific

	databases, including EU repositories that can prompt towards finding CALIMERO's information. Otherwise, since data will also be included in internal repositories (accessible e.g., through the project website), search options by keywords and filters might be included to facilitate and speed up the access to data.
Identifiability of data	Data will be registered according to metadata descriptive rules, as illustrated in the "metadata usage" documentation. Some identification mechanisms such as Digital Object Identifiers (DOIs) will also be considered in cases of data published as part of journal or book articles.
Versioning	Data updates will be performed during the project if required by the task activities (e.g., in the case of specific methodological constraints or if requested by peer-reviewers). In general, however, data will not undergo updating processes (i.e., versioning) starting from the end of the project, unless additional funding or project follow-up will not be foreseen. The developed or improved LCSA for the different functions of impact assessment and calculation of indicators in WP3 will be conceived in a way that the methodological framework will allow to perform data and metadata updates by any user. To this aim, data sources will be consistently documented in every data repository of the WP3.
Metadata usage	Several standards are available to set up metadata. Depending on the type of generated data and datasets, the WP3 team will consider to use one or more of the following standards: • Data Documentation Initiative (DDI), which is a metadata standard produced by the DDI alliance and used mainly in the fields of soft sciences (social, behavioural, economy, health science) to describe "surveys and other observational methods" (https://ddialliance.org/); • Ecological Metadata Language (EML), which is a metadata standard in the fields of environmental sciences (https://eml.ecoinformatics.org/); • INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119 (https://inspire.ec.europa.eu/auto-created-vocabulary-2/guidance-document); • ISO/IEC 11179-1:2023, Information technology — Metadata registries (MDR) — Part 1: Framework (https://www.iso.org/standard/78914.html) (not exhaustive list)
2.2 Making data openly accessible	
Accessibility	All data generated within the WP3 will be disclosed and made open access through deliverables and scientific/technical reporting, such as peer-reviewed journal publications.
Method of availability	Data will be made available in a shared repository of the project consortium (accessible e.g., through the project website), and further included in proprietary but open access repositories of the journals where the articles containing the data will be published. Moreover, according to Horizon Europe's rules, metadata will be published under a public domain licence (CC0 or equivalent) or a very open licence (CC-BY or equivalent), allowing the reuse of the metadata without restrictions (or with very limited restrictions for CC-BY).

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Methods/software needed to access the data	No specific methods or software will be needed to access the data and metadata of WP3.
Access control	Access to data will not be restricted, and thus no access control is needed.
2.3 Making data interoperable	
Interoperability	To be defined.
2.4 Data reuse and quality	
Licensing	Data will not be released under any specific licence. However, data incorporated in published articles will rely on the specific licence of use foreseen by the publication journal, book or scientific/technical frame.
Data availability and embargo	Data generated in each task will be available at the time of release of the task deliverable(s). An embargo period for some deliverables (kept confidential) might be foreseen for those data incorporated in manuscripts submitted to journals during the peer-review process.
Reuse restrictions	No restrictions on the reuse of the data are foreseen, as long as the licence rules under which some data may belong are respected (e.g., CC0 or equivalent, or CC-BY or equivalent).
Data retention	Data stored in the project internal repositories will be kept accessible for 2 years, while those included in published documents will stay for an indefinite time. Data included in open-source platforms will be available in an indefinite time.
Data quality	To be defined. In principle, Data Quality Assessment (DQA) procedures as typically applied in LCA might be considered (e.g., use of pedigree matrix approach, where the reliability of data source and the conformance of the dataset are analysed with expert-based/Likert-type based scoring systems).
3 Other research outputs	
3.1. Plan for the management of other research outputs	Not applicable.
3.2. FAIR outputs	Not applicable.
3 Allocation of resources	
3.1 Costs	To be estimated. In principle there will not be costs additional to those already foreseen within the CALIMERO budget assigned to each beneficiary partner.
3.2 Data management responsibilities	Data management responsibility is to be attributed to the owner/producer of the data.
3.3 Costs of preservation	To be estimated.
4 Data security	
4.1 Data security	Data will be open access, and therefore security rules will belong to the type of licence that is assigned to the data.
5 Ethical aspects	
5.1 Ethical aspects	No ethical aspects will concern the elaboration and management of environmental or economic data.
6 Other aspects	
6.1 Other aspects	Not applicable at this stage.

1.4 WP4. FRAMEWORK FOR THE SUSTAINABLE OPTIMIZATION OF TARGET BIO-BASED SECTORS

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Work package	WP4. FRAMEWORK FOR THE SUSTAINABLE OPTIMIZATION OF TARGET BIO-BASED SECTORS
1.Data summary	
1.1. Purpose of data collection/generation	Development and validation of multi-objective optimization framework suitable for bio-based sectors, including life cycle sustainability assessment models, simulation models potentially improving the sustainability performance of industrial processes and uncertainty analysis.
1.2. Relation to project objectives	The data collected and generated in WP4 will significantly contribute to: • O4 - To develop a MOO framework of industrial processes that integrates the improved LCSA methodologies based on PEF, for optimization of bio-based industrial processes with sustainability indicators • O5 – To provide feasible solutions with better sustainability performance than the current situation for the five studied bio-based industry sectors and the procedures to monitor their sustainability performance
1.3. Types/format of data	Type: all data will be considered confidential data ✓ Graphical data: Flow diagrams ✓ Modelling: simulation, LCSA models already available ✓ Structured data: life cycle inventories, mass balances ✓ Numerical data: LCSA results, uncertainty analysis results ✓ Multi-objective optimization (MOO): algorithm ✓ Text: interpretation of results Format: ✓ Reports: .docx, pdf ✓ Tables: .xlsx, .csv ✓ Figures: jpeg, .png ✓ LCSA results: csv, .docx ✓ Models: .bkp. .py, MAT ✓ MOO: .py
1.4 Origin of data or reuse of existing data	In this WP most of the data will be reused and adapted to fit in Calimero's objectives, given the aim of developing and validating the MOO framework to be used alter in WP5. The only exception is task 4.3, where the simulation of potential solutions improving sustainability performance of industrial processes will be modelled, using existing methods, but using the processes modifications suggested in WP2. Potential solutions: WP2 and WP5 ✓ LCSA methodology: WP3 ✓ LCSA background data: existing LCA databases ✓ Inventory: WP2 and literature review ✓ MOO: LIST and CTA will use their own developed methods and adapt it to bio-based industries needs and LCSA methodological development done in WP3. ✓ Simulation models: methods used by DTU and found in literature will be adapted and sued for the specific processes to be assessed in WP4
1.5 Scale of data	Generated data: 2 MOO algorithms and 7-12 simulations <50Gb

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1.6 Data utility	CTA, LIST, DTU will use it in WP5 Industrial partners will benefit of the results obtained later in WP5 by using the outcomes from WP4.
2. FAIR Data	
2.1. Making data findable, including provision for metadata	
Facilitating findability	Life Cycle Assessment, Process simulation, industrial modelling, sustainability, multi-objective optimization, artificial intelligence
Identifiability of data	Most data will remain confidential, but some publications can result after using the outcomes of WP4 and will do have a DOI.
Versioning	Versioning will be used.
Metadata usage	No metadata format is expected to be used
2.2 Making data openly accessible	
Accessibility	Industrial data and MOO algorithms will remain confidential due to exploitation issues. Conclusions and main methodology used in WP4 will be shared in a public deliverable
Method of availability	Public deliverables will be shared in the website of Calimero
Methods/software needed to access the data	Internet browser and internet connection
Access control	No restriction of access
2.3 Making data interoperable	
Interoperability	Data will be published to ensure they are findable, accessible, and interoperable and reusable internally (not by public), given the confidentiality of MOO algorithms developed
2.4 Data reuse and quality	
Licensing	No
Data availability and embargo	No
Reuse restrictions	not applicable
Data retention	At least 5 years after the end of the project
Data quality	Uncertainty analysis and validation procedures will be applied to test the performance of the models developed
3 Other research outputs	
3.1. Plan for the management of other research outputs	The outcomes and conclusions from WP4 will be considered in the final guidelines and recommendations published in WP5
3.2. FAIR outputs	Some outcomes will remain confidential thus not will be reusable by external parties. However, the methodologies will be made findable, accessible and interoperable.
3 Allocation of resources	
3.1 Costs	Costs are included in project budget
3.2 Data management responsibilities	Task leaders will be responsible of the data collected to perform the work of each specific task
3.3 Costs of preservation	Costs are included in project budget. Long-term storage of data costs will be assumed by each partner.
4 Data security	
4.1 Data security	Data will be stored in online servers with backup, recovery and secure storage services.
5 Ethical aspects	
5.1 Ethical aspects	NA

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

6 Other aspects	
6.1 Other aspects	NA

1.5 WP5. IMPROVEMENT OF THE SUSTAINABILITY PERFORMANCE OF BIO-BASED INDUSTRIES

Work package	WP5. IMPROVEMENT OF THE SUSTAINABILITY PERFORMANCE OF BIO-BASED INDUSTRIES
1. Data summary	
1.1. Purpose of data collection/generation	Identification of potential industrial solutions, sectorial and cross-sectorial, to improve sustainability performance of bio-based industries and demonstrate their feasibility using the MOO framework from WP4.
1.2. Relation to project objectives	O5 - To provide feasible solutions with better sustainability performance than the current situation for the five studied bio-based industry sectors and the procedures to monitor their sustainability performance Contribution to O5 by identifying and analysing industrial solutions O6- To maximize the impact of CALIMERO through tailored Exploitation, Dissemination and Communication activities aiming to pave the way to market of the sustainability assessment tool, build synergies with other R&D projects and transfer the project results to different target audiences Contribution to O6 by preparing guidelines and recommendations for different stakeholders
1.3. Types/format of data	Type: all data will be considered confidential data except guidelines and recommendations ✓ Graphical data: Flow diagrams ✓ Modelling: simulation, LCSA models already available ✓ Structured data: life cycle inventories, mass balances ✓ Numerical data: LCSA results, uncertainty analysis results ✓ Multi-objective optimization (MOO): algorithm ✓ Text: interpretation of results Format: ✓ Reports: .docx, pdf ✓ Tables: .xlsx, .csv ✓ Figures: jpeg, .png ✓ LCSA results: csv, .docx ✓ Models: .bkp. .py, MAT ✓ MOO: .py
1.4 Origin of data or reuse of existing data	The MOO will be fed from WP4. All the consortium will contribute to identify sectorial and cross-sectorial solutions at plant-level. Best Available Technologies will also be sued as source, as well as relevant literature available.
1.5 Scale of data	<5Gb
1.6 Data utility	Industrial actors: they will obtain strategies to optimize their sustainability performance with technical and sustainability criteria. Other stakeholders like policy makers, consumers or researches, will obtain guidelines and recommendations on how to perform LCSA to improve the sustainability performance of bio-based industries.
2. FAIR Data	
2.1. Making data findable, including provision for metadata	

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Facilitating findability	Life Cycle Assessment, Process simulation, industrial modelling, sustainability, multi-objective optimization, artificial intelligence, recommendations, policy-making
Identifiability of data	Yes
Versioning	No
Metadata usage	Not applicable
2.2 Making data openly accessible	
Accessibility	Guidelines and recommendations will be publicly available. Sustainability performance improvements and solutions for industries will remain confidential due to exploitation issues
Method of availability	Calimero website, ZENODO repository, CORDIS
Methods/software needed to access the data	Internet browser and internet
Access control	No access restrictions to public deliverables, guidelines and policy recommendations
2.3 Making data interoperable	
Interoperability	Guidelines and recommendations to adapt PEF method to bio-based industries will be reported using similar vocabulary and structure than PEF method report.
2.4 Data reuse and quality	
Licensing	No
Data availability and embargo	The LCSA methodology will be available to be used by any LCSA practitioner
Reuse restrictions	The LCSA and MOO models developed will not be publicly shared
Data retention	At least 5 years after the end of the project
Data quality	Internal review by different experts in the consortium will be performed before publishing any date
3 Other research outputs	
3.1. Plan for the management of other research outputs	Not applicable
3.2. FAIR outputs	Some outcomes, like LCI inventories or simulation models will remain confidential thus not will be reusable by external parties. However, the methodologies will be made findable, accessible and interoperable.
3 Allocation of resources	
3.1 Costs	Costs are included in project budget
3.2 Data management responsibilities	Task leaders will be responsible of the data collected to perform the work of each specific task
3.3 Costs of preservation	Costs are included in project budget. Long-term storage of data costs will be assumed by each partner.
4 Data security	
4.1 Data security	Data will be stored in online servers with backup, recovery and secure storage services.
5 Ethical aspects	
5.1 Ethical aspects	NA
6 Other aspects	
6.1 Other aspects	NA

1.6 WP6. EXPLOITATION, DISSEMINATION AND COMMUNICATION

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Work package	WP1. EXPLOITATION, DISSEMINATION AND COMMUNICATION
1.Data summary	
1.1. Purpose of data collection/generation	To share information. To raise public and scientific awareness about the outcomes of the project and the developments achieved and to maximise the impact of the project's results through appropriate exploitation strategies
1.2. Relation to project objectives	To ensure that the project's objectives are widely promoted to the target groups defined on a European level and beyond through an appropriate communication strategy.
1.3. Types/format of data	The project website, press releases, brochures, business models, exploitation plan, posters, presentations. text files, multimedia, .pdf, .pptx, .odf, .xls, .mp3, .mp4
1.4 Origin of data or reuse of existing data	Data is provided by project members and generated within the project.
1.5 Scale of data	<1 TB
1.6 Data utility	The public in general, researchers, research communities, decision-makers, industry.
2. FAIR Data	
2.1. Making data findable, including provision for metadata	
Facilitating findability	Search keywords will be provided to search for and successfully find WP6 outputs related to CALIMERO projects and its promotion.
Identifiability of data	No
Versioning	Yes
Metadata usage	No metadata standards will be used.
2.2 Making data openly accessible	
Accessibility	Outputs will be made openly available. All data that must be public and does not affect IPR issues will be shared and openly available.
Method of availability	Online
Methods/software needed to access the data	Web browser for web content, open-source tools for documents.
Access control	no
2.3 Making data interoperable	
Interoperability	NA
2.4 Data reuse and quality	
Licensing	Data will be licensed using standard licences (i.e., Creative Commons licences) in line with the obligations set out in the Grant Agreement.
Data availability and embargo	Data will be available straight after publication.
Reuse restrictions	Permissions are provided through licences.
Data retention	At least 5 years after the project ends.
Data quality	Multiple authors involved in data creation and internal revision before publishing
3 Other research outputs	
3.1. Plan for the management of other research outputs	If the research outputs are not protected (under private IP means) they will be shared through the website and social networks.
3.2. FAIR outputs	Research outputs from other WPs will be made open and free available through scientific publications that will be found in the webpage, and repositories such as Zenodo, Openair and so on.

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

3 Allocation of resources	
3.1 Costs	All costs for making data FAIR are integrated within the project
3.2 Data management responsibilities	Task leaders will be responsible of the data collected to perform the work of each specific task
3.3 Costs of preservation	Costs are included in the project budget. Long-term storage of data costs will be assumed by each partner.
4 Data security	
4.1 Data security	Data will be stored in online servers with backup, recovery and secure storage services.
5 Ethical aspects	
5.1 Ethical aspects	NA
6 Other aspects	
6.1 Other aspects	No

1.7 WP7. PROJECT MANAGEMENT

Work package	**WP7. PROJECT MANAGEMENT**
2. Data summary	
1.1. Purpose of data collection/generation	Data is collected within WP7 to obtain information, to share information, to keep on record, to combine with other data and to make informed decisions to fulfil the objectives of the project as defined by the Grant Agreement.
1.2. Relation to project objectives	To ensure effective project management and appropriate scientific coordination
1.3. Types/format of data	1. Project documentation (procedures, plans, metrics, risks, meetings, presentations) 2. Deliverables and milestones 3. Efforts and financial data The format of the data: plain text, .pdf, .docx, .pptx, .odf, .xls
1.4 Origin of data or reuse of existing data	Data is produced and provided by project members
1.5 Scale of data	<500GB
1.6 Data utility	The target groups for data generated in WP1 include the project officer, the project consortium/members, the EC, external researchers and research communities, industry, decision-makers, and the public in general
2. FAIR Data	
2.1. Making data findable, including provision for metadata	
Facilitating findability	Search keywords will be provided to search for and successfully find WP7 outputs.
Identifiability of data	No
Versioning	Yes. Versions are frequently monitored to discard those that are not required for verification, reproducibility, or transparency, amongst others
Metadata usage	No metadata standards are expected to be used
2.2 Making data openly accessible	
Accessibility	1. Shared with the consortium to support project work. Some datasets will be shared under restricted access conditions via the Private Area repository in the website.

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

	2. All deliverables and milestones are shared within the consortium and with the EC. Public deliverables will be deposited in the project website for long term preservation and curation. 3. Shared with the consortium and EC. Contains personal/sensitive/confidential financial data so is kept private.
Method of availability	Via the project website
Methods/software needed to access the data	Web browser for web content, open-source tools for documents.
Access control	Via open-source methods of authorisation and authentication. Data access controls, such as passwords or firewalls, will be used to limit access to confidential data and protect it from unauthorised changes.
2.3 Making data interoperable	
Interoperability	Via the use of well-known formats such as those used by Microsoft Word and Excel
2.4 Data reuse and quality	
Licensing	NA
Data availability and embargo	NA
Reuse restrictions	No
Data retention	At least 5 years after the project ends.
Data quality	No
3 Other research outputs	
3.1. Plan for the management of other research outputs	Not applicable as research outputs will be managed within each WP and WP6.
3.2. FAIR outputs	NA
3 Allocation of resources	
3.1 Costs	All costs for making data FAIR are integrated within the project
3.2 Data management responsibilities	Task leaders will be responsible of the data collected to perform the work of each specific task
3.3 Costs of preservation	Long term preservation of data will be ensured by the project and by the partners themselves.
4 Data security	
4.1 Data security	Data will be stored in online servers with backup, recovery and secure storage services.
5 Ethical aspects	
5.1 Ethical aspects	NA
6 Other aspects	
6.1 Other aspects	No

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 European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.