



TEALHELIX

SMART LABELS
INCLUSIVE SOLUTIONS

Ref. Ares(2025)1299523 - 18/02/2025

DELIVERABLE 1.2

Data Management Plan 1

KU Leuven

February 2025



PROJECT INFORMATION

PROGRAMME	Horizon Europe
TOPIC	HORIZON-CL6-2023-COMMUNITIES-01-6
TYPE OF ACTION	HORIZON Research and Innovation Actions
PROJECT NUMBER	101136955
START DAY	1 September 2024
DURATION	48 months

DOCUMENT INFORMATION

TITLE	DATA MANAGEMENT PLAN 1
WORK PACKAGE	1
TASK	Task 1.4 Data Management
AUTHORS (Organisation)	Florian Lange (KUL)
REVIEWERS	Justina Barsyte (KUL), Jan Michael Bauer (CBS)
DATE	28/02/2025

DISSEMINATION LEVEL

PU	Public, fully open	x
SEN	Sensitive, limited under the conditions of the Grant Agreement	
Classified R-UE/EU-R	EU RESTRICTED under the Commission Decision No2015/444	
Classified C-UE/EU-C	EU CONFIDENTIAL under the Commission Decision No2015/444	
Classified S-UE/EU-S	EU SECRET under the Commission Decision No2015/444	

DOCUMENT HISTORY

VERSION	DATE	CHANGES	RESPONSIBLE PARTNER
V0.1	30/01/2025	The first draft of DMP V1 is ready for internal review	KUL
V0.2	06/02/2025	Reviewers' feedback	KUL/CBS
V1.0	11/02/2025	Revision of DMP V1 taking into account reviewers' feedback	KUL

LEGAL NOTICE

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.

© TEALHELIX Consortium, 2025

Reproduction is authorised provided the source is acknowledged.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	8
1. ABOUT TEALHELIX	8
1.1 Project abstract.....	8
1.2 Consortium	8
2. DATA SUMMARY	9
3. FAIR DATA.....	15
3.1 Making data findable, including provision for metadata	15
3.2 Making data accessible.....	15
3.2.1 Repository	15
3.2.2 Data.....	16
3.2.3 Metadata	17
3.3 Making data interoperable	17
3.4 Increase data re-use	18
4. OTHER RESEARCH OUTPUTS.....	19
5. ALLOCATION OF RESOURCES.....	19
6. DATA SECURITY	20
7. ETHICS	21
8. OTHER ISSUES	22

LIST OF TABLES

Table 1: TealHelix Consortium 8

Table 2: Data to be generated or used through TealHelix activities 11

ABBREVIATIONS

AI	Artificial Intelligence
D	Deliverable
DMP	Data Management Plan
DOI	Digital Object Identifier (DOI)
OSF	Open Science Framework
TBD	To be determined

Executive Summary

The Data Management Plan (DMP) for the Horizon Europe project TealHelix is designed to ensure proper, high-quality management of the data generated during the project. It describes the data generated during TealHelix project activities, the potential for data exploitation beyond the end of the project, and the overarching data management strategy.

The DMP follows the template for “EU Grants: Data management plan (HE): V1.1 – 01.04.2022” and focuses on the implementation of the FAIR principles for making research data findable, accessible, interoperable and reusable. It also describes how open science practices are implemented as a cornerstone of TealHelix data management to maximize the reliability, utility, and reuse potential of the data and other research outputs to be generated. The DMP specifies how TealHelix data will be handled in a secure and ethically responsible way in line with the European Code of Conduct for Research Integrity and the General Data Protection Regulation.

The DMP is aligned with the TealHelix Project Management Plan (Deliverable 1.1). Its implementation will be guided by the project’s Data Manager from KU Leuven. The DMP will be updated as additional details about the data to be generated during TealHelix activities become known, and updated versions of the DMP will be published in months 18, 36, and 48.

1. ABOUT TEALHELIX

1.1 Project abstract

TealHelix advances the state-of-the-art by proposing a more precise and targeted approach to sustainable food labelling – empowerment through personalization and inclusion. Using the underlying logic of motivational matching, we will develop several new labelling approaches and digital social innovations to guide and improve consumer decision-making. Such an approach will also enable us to address resistance to sustainability ideas and tailor our interventions to the heterogeneity in individual needs of vulnerable consumers. Combining insights from life cycle, social and economic environments analysis, measurement, and consumer behaviour theories, we will develop a new measure to assess how individual and planetary preferences for various sustainability dimensions can be aligned to reach sustainability goals. Next, we will test a number of means of transmission: traditional labelling approaches, digital and brick-and-mortar retail labelling approaches, and smart labelling approaches. To sustain and scale the change, we will develop integrity guidelines and new sustainability information provision standards for the industry. We will leverage the interdisciplinary and transdisciplinary competencies of the consortium in marketing, consumer behaviour, psychology, environmental and information sciences, as well as in communication, retailing, and standard-setting industries. The project will generate multiple novel methods to study labelling approaches and original empirical evidence through machine learning-based 'big data' analysis, large-scale surveys, experience-sampling, and micro-level experiments. Finally, we will integrate the findings into digital social innovations powered by AI tools to support labelling solutions. As a result, we will provide a deeper understanding of how various external environments shape attitudes and beliefs towards food sustainability labelling, how to motivate consumers to use and follow sustainable label information, and how to include the ones who are in greatest need.

1.2 Consortium

TealHelix leverages the interdisciplinary and transdisciplinary competencies of the consortium in marketing, consumer behaviour, psychology, environmental and information sciences, as well as in communication, retailing, and standard-setting industries.

TealHelix is implemented by a well-balanced consortium comprised of high-level researchers in the areas of marketing, consumer behaviour, psychology, environmental and information sciences as well as industry experts. Our 18 partners represent a total of 11 different countries (Austria, Belgium, Denmark, Germany, Greece, Estonia, Latvia, Lithuania, the Netherlands, Poland, and Switzerland). The consortium is a high-quality partnership based on the diversified expertise, knowledge, and networks substantially covering all skills, resources, and access to stakeholders required for successful implementation.

Table 1: TealHelix consortium

No	Legal name	Short name	Country
1	KATHOLIEKE UNIVERSITEIT LEUVEN	KUL	BE
2	SAFE FOOD ADVOCACY EUROPE	SAFE	BE
3	WHITE RESEARCH SRL	WR	BE

4	VILNIAUS UNIVERSITETAS	VilniusU	LT
5	AdCogito Elgsenos tyrimu institutas Vsl	AdC	LT
6	VALSTYBINE MAISTO IR VETERINARIJOS TARNYBA	VMVT	LT
7	SIA "Rimi Baltic"	RIMI	LV
7a	RIMI ESTONIA	RIMI EE	EE
8	ZENITH POLAND SP Z O.O.	Zenith	PL
9	Stowarzyszenie	SAR	PL
10	UNIVERSITY OF MACEDONIA	UoM	EL
11	EREVNITIKO PANEPISTIMIAKO INSTITOUTO SYSTIMATON EPIKOINONION KAI YPOLOGISTON	ICCS	EL
12	RIJKSUNIVERSITEIT GRONINGEN	RUG	NL
13	STICHTING VU	VU Amsterdam	NL
14	GS1 GERMANY GMBH	GS1	DE
15	UNIVERSITAET FUER BODENKULTUR WIEN	BOKU	AT
16	COPENHAGEN BUSINESS SCHOOL	CBS	DK
17	UNIVERSITAT ST GALLEN	USG	CH

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.

Some tasks will make use of existing datasets, for example, to characterize the environmental, social, and economic impact of food products (Task 3.2 and 3.3) or scraping sustainability-related data from online platforms (Task 2.4). Other tasks will be informed by literature reviews of prior research, but address research questions that require the collection of new data.

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

TealHelix combines qualitative and quantitative research methods.

Qualitative studies (e.g., interviews, focus groups) will involve the generation of audio recordings and of transcripts (i.e., text documents) of those recordings. Transcripts will be prepared in the language that was used during the interview/focus group (which may be English or the local language). The partner conducting the qualitative research will also prepare a summary of the results (in text format) in English. Text data will be published in openly accessible file formats (e.g., Plain text, Unicode,

pdf). In addition, text data may be published in additional, software-specific file formats (e.g., NVIVO) where this practice can be expected to facilitate reuse of the data.

Quantitative studies (e.g., surveys, experiments, life cycle analyses, data scraping) will produce data in spreadsheet (i.e., tabular) form. Spreadsheets will generally pertain to how one or multiple variables (e.g., responses to observed events, experimental stimuli, or survey questions) vary across observation units, and they will allow for quantitative analysis. Spreadsheet data will be published in openly accessible file formats (e.g., csv, pdf). In addition, spreadsheet data may be published in additional, software-specific file formats (e.g., SPSS, Stata) where this practice can be expected to facilitate reuse of the data.

Details about the types and format of the data to be collected and analysed within the TealHelix tasks can be found in Table 2. This information will be updated in future versions of the DMP.

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

All data generation and re-use activities allow addressing research questions and/or evaluating innovation activities related to the TealHelix project objectives:

- O1. Analyse the influence of media/marketing and sociocultural aspects on consumer understanding of sustainability.
- O2. Identify information expectations and needs related to all three dimensions of sustainability.
- O3. Develop a measure for consumers that matches preferred and the most effective sustainability dimensions.
- O4. Develop novel behavioural interventions to boost the motivation to use sustainability information.
- O5. Engage citizens in innovation creation through co-creation and citizen science.
- O6. Develop AI based apps for citizens.
- O7. Pilot test the effectiveness of different labelling approaches.
- O8. Build capabilities, standards, and trust in sustainable food labelling schemes.
- O9. Deliver practical tools and guidelines for national and regional authorities in the EU and Associated Countries.

Table 2 details the specific objectives related to task-specific data generation and re-use activities.

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

Table 2 shows how the size of the dataset to be generated will differ across TealHelix tasks.

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

Data will originate from existing data bases, from qualitative and quantitative research activities, and from the evaluation of trainings focused on stakeholder capacity building (see Table 2).

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

Data-based reports will be useful to a variety of stakeholders within and beyond academia (e.g., advertisers, retailers, consumers, consumer rights organizations, policy makers). The datasets we will generate will primarily be of interest to other researchers. They will be optimized for secondary research use before being published.

Table 2: Data to be generated or used through TealHelix activities

Task	Data type	Data format	Purpose	Size	Origin	Data utility beyond the project
2.1	inventory of sustainability labelling approaches	spreadsheet	providing systematic data about the key used sustainable food labelling approaches	TBD	literature review	policymakers, government agencies, businesses, industry leaders, academics, researchers, media firms, European organisations
2.1	responses to interview questions about the challenges, perceptions, and effectiveness of current sustainable food labelling practices	text document (transcripts)	understanding perceptions of and engagement with labelling approaches	TBD	interviews with key actors of the pilot countries	policymakers, government agencies, businesses, industry leaders, academics, researchers, media firms, European organisations
2.2	survey responses related to demographics, psychological traits, preferences, and behaviour for sustainability food labelling	spreadsheet	exploring external factors influencing consumer behavior and segmentation based on sustainability preferences and demographics	6-10MB	survey in multiple EU countries	policymakers, government agencies, businesses, industry leaders, academics, researchers, media firms, European organisations
2.3	presence, share, and effectiveness of communicated sustainability claims	spreadsheet, possibly images/videos	to capture the state of sustainability shown in marketing communication	TBD	media monitoring across multiple countries	brands, media and creative agencies, sustainability/c communication/marketing scientists
2.3	responses to focus group questions about sustainability in communication	text document (transcripts), audio recordings	to capture and understand the challenges of making sustainability part of the brands' communication	TBD	focus groups with advertising experts	brands, media and creative agencies, sustainability/c communication/marketing scientists

2.4	observation records of natural behaviours, voiced opinions, and emotions in online settings (e.g., social media posts and their metadata)	spreadsheet , possibly Images	analysis of opinions, emotions, conversation topics, sentiment etc.	~200 MB, larger if visual data are included	data scraping from online platforms	other behavioural scientists
3.1	responses to experience sampling questions about consumers' information needs and information search in supermarkets and about potential contextual influence factors	spreadsheet	understanding consumers' information need, information search methods, shopping behaviours and contextual influence factors	1-100 MB	experience sampling study	other behavioural scientists, training target groups (i.e., producers, retailers, and public officials), food apps developers
3.1	choices, response times, stated and revealed preferences about food and sustainability-related concepts and labels	spreadsheet	understanding consumers' implicit associations between food products and different sustainability dimensions	1-100 MB	laboratory/ online experiments	other behavioural scientists, training target groups (i.e., producers, retailers, and public officials)
3.2	product-level environmental impact information	spreadsheet	assessing the environmental impact of different food products through life cycle analysis	1-100 MB	life cycle data from existing databases and from primary data, life cycle analysis	other scientists working on LCA and environmental impacts of food products, policy makers, industries
3.3	product-level social and economic impact information	spreadsheet	assessing the social and economic impact of different food products through nutri-score, animal welfare and economic data	1-100 MB	social and economic assessment data from existing data bases, social and economic assessment	other scientists working on socio-economic sustainability impacts of food products, policy makers, industries
3.4	items to assess consumers' relevance of environmental, social, and economic sustainability dimensions	text document	capture consumer perspective as part of the Sustainable Food Compass	TBD	systematic literature review, focus groups and in-depth interviews	the data will lead to the development of a scale that will be useful for future research
3.4	survey responses related to the sustainable food compass and consumer characteristics	spreadsheet	understand consumer relevance of food sustainability aspects and differences across consumer cohorts	1-10MB	survey in multiple countries	data might be used for additional analyses by researchers

4.1	experimental data (choices, response times, accuracy in response to knowledge questions, stated and revealed preferences about food and sustainability-related concepts and labels, survey responses related to psychological characteristics)	spreadsheet	testing interventions to promote sustainability-related outcomes	1-10 MB	laboratory/ online <i>experiments</i>	other behavioural scientists and sustainable food retailers
4.2	experimental data (see Task 4.1)	spreadsheet	testing interventions to promote sustainability-related outcomes	1-10 MB	laboratory/ online experiments	other behavioural scientists and sustainable food retailers
4.3	experimental data (see Task 4.1)	spreadsheet	testing interventions to promote sustainability-related outcomes	1-10 MB	laboratory/ online experiments	other behavioural scientists and sustainable food retailers
4.4	experimental data (see Task 4.1)	spreadsheet	testing interventions to promote sustainability-related outcomes for vulnerable consumers	1-10 MB	laboratory/ online experiments	other behavioural scientists and sustainable food retailers
5.1	responses & interactions of stakeholders taking part in co-creation sessions reflecting on various aspects of the project	text document (transcripts; meeting reports), audio recording, online meeting recording, spreadsheet	co-creation: gathering input from various stakeholders to improve project output	1-10GB (strongly depends on whether sessions will be online and recorded or not)	co-creation sessions	other behavioural scientists, as well as the training target groups (i.e., marketing and communication professionals, producers, retailers, and policymakers)
5.2	aggregated data from prior tasks: consumers choices, response times, accuracy in response to knowledge questions, stated and revealed preferences about food and sustainability-related concepts and labels, survey responses related to psychological characteristics, demographics	spreadsheet	integrating prior findings into the guidelines for pilot testing	1-10MB	data collected in other tasks	other behavioural scientists, as well as the training target groups (i.e., producers, retailers, and public officials)

6.1	list of interventions that are likely to be effective	text document	gathering practical insights to enhance the effectiveness of behavioural interventions	1-10MB	expert interviews/ focus groups	other behavioural scientists, marketers, advertisers, sustainable food producers and retailers
6.1	choices, response times, accuracy in response to knowledge questions, stated and revealed preferences about food and sustainability-related concepts and labels, survey responses related to psychological characteristics	spreadsheet	testing means o presentation of sustainability labels and related information	1-100MB	laboratory/ online/ field experiments	other behavioural scientists, marketers, advertisers, sustainable food producers and retailers
6.2	observation records of purchases and app use, survey responses related to psychological characteristics	spreadsheet	testing means to transmit sustainability-relevant information and interventions in the pilots	1-100 MB	field experiments in physical and online stores	other behavioural scientists, retailers, (food) app developers
6.3	observation records of app use	spreadsheet	optimizing TealHelix tools	TBD	Task 6.1, Task 6.2	TBD
7.2	evaluation survey responses related to knowledge about integrity aspects of food sustainability information provision	spreadsheet	test if professionals (marketing and communication professionals, producers, retailers and policymakers) have the needed knowledge and capacity to select and apply ethical labelling schemes	1-100 MB	training session evaluation survey	other behavioural scientists, as well as the training target groups, i.e., marketing and communication professionals, producers, retailers, and policymakers
7.3	evaluation survey responses related to knowledge about sustainability food promotion	spreadsheet	test if professionals (producers, retailers, and public officials) know how to efficiently promote sustainable food	1-100 MB	training session evaluation survey	other behavioural scientists, as well as the training target groups, i.e., producers, retailers, and public officials

3. FAIR DATA

TealHelix partners will deposit all datasets that are suitable for publication in an open-science repository (Zenodo or OSF) that supports the FAIR principles through assignment of persistent identifiers, creation of metadata, and provision of standard open-access reuse licences. Reports and publications based on those datasets will include links to the deposited datasets. Datasets will not be published directly on the project website, but the deposited datasets will link to the project website. Datasets will be published as soon as possible and ideally before the end of the Work Package that generated the data, unless it is in the partners' legitimate interest to keep the data under embargo until the publication of research output (e.g., a scientific article) that is based on the respective dataset.

3.1 Making data findable, including provision for metadata

Will data be identified by a persistent identifier?

Each published dataset will receive a Digital Object Identifier (DOI) at Zenodo or the OSF.

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.

Metadata will be created for each published dataset. Metadata standards have not been established in all of the scientific disciplines represented in TealHelix. For each published dataset, metadata will include at least the following information: title, description (including grant information), keywords, language, creation date, modification date (if any). In addition, each published dataset will be accompanied by documentation (e.g., a codebook) that ensures unambiguous understanding of variable names, labels, and values.

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

Metadata will include keywords that support the findability of published datasets.

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

At both Zenodo and the OSF, metadata of datasets with DOI are indexed and searchable via repository search engines.

3.2 Making data accessible

3.2.1 Repository

Will the data be deposited in a trusted repository?

Both Zenodo (dedicated repository for projects funded by the European Commission) and the OSF are generally trusted repositories.

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

Given that publishing data and metadata is not linked to direct expenses at Zenodo and the OSF, and all TealHelix partners have the opportunity to share datasets via those channels, we have not explored project-specific arrangements with the repositories.

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

At both Zenodo and the OSF, datasets can be linked to DOIs and DOIs resolve to the published dataset.

3.2.2 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.

As a rule, all datasets will be made openly available. Personal data within datasets will be made unrecognizable before they are shared, for example, by anonymizing the data. Project partners will take appropriate measures to avoid disclosure of research participants' identity and to protect research participants' personal data. If, in exceptional cases (e.g., interview recordings), data cannot be anonymized for open sharing, these data will be kept on file in secure storage and access to these data can be given to researchers upon proper ethical and legal clearance for the duration of the retention period. In such exceptional cases, TealHelix partners should make reasonable efforts to share datasets that are as rich as possible while ensuring participants' anonymity (e.g., partially redacted interview transcriptions). If, in other exceptional cases, it goes against the legitimate interest of a project partner to openly share a full dataset, the partner should make reasonable efforts to share a dataset that is as rich as possible while ensuring their legitimate (i.e., business) interest.

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.

As a rule, datasets will be made openly available as soon as they are ready for publication. When it is in a partner's legitimate interest to keep the data under embargo until the publication of research output (e.g., a scientific article), the publication of a dataset can be delayed until the publication of the research output that is based on the respective dataset.

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

All data will be published on Zenodo or the OSF and, as such, they will be freely available for download via https.

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

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

As a rule, there are no restriction on use. Data that cannot be made openly available (e.g., interview recordings) will be kept on file in secure storage and access to these data can be given to researchers upon proper ethical and legal clearance for the duration of the retention period. Access to these data will be managed by the ethical and legal authority relevant at the institution of the partner who stores the data.

3.2.3 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 records will be openly available and CC0 licensed. Metadata for datasets with restricted access (if any) will include information on how access to the dataset can be requested.

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

Anonymized datasets and metadata will be made permanently available on the open-science repositories. Hence, they will be retained for the lifetime of the respective repository.

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)?

All published datasets will be made available in openly accessible file formats. In addition, data records may include the datasets in software-specific file formats (e.g., SPSS, Stata) where this practice can be expected to facilitate reuse of the data. Where available, processing and analysis scripts or code will be published alongside published datasets.

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 a highly interdisciplinary project, TealHelix does not follow the standard data sharing procedures of any particular discipline, and in many of the involved disciplines, such standard procedures have not been established yet. Project partners will ensure the interoperability of datasets by following common procedures in their respective fields. This may include 1) using openly accessible standard file formats, 2) providing data in unprocessed and processed form, 3) supplementing datasets with processing and analysis scripts, 4) the use of standard table formats and consistent formatting of cell values (in the case of spreadsheet data), and 5) supplementing datasets with documentation (e.g., codebooks) that ensures unambiguous understanding of variable names, labels, and values.

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?

Due to the lack of commonly applied data and metadata ontologies in the disciplines contributing to TealHelix, mapping TealHelix data and metadata to such ontologies is not feasible. Published datasets will be accompanied by openly available documentation that ensures unambiguous understanding of variable names, labels, and values.

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

Where applicable, metadata will include qualified references to the persistent identifiers of additional datasets that complement the dataset at hand. These references will specify the type of relation between the datasets.

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.)?

Along with each published TealHelix dataset, corresponding documentation will be published in the same trusted repository. This documentation will include a file (in openly accessible format) that explains any variable names, labels, and values included in the accompanying dataset (e.g., a codebook). Where applicable, processed datasets will also be accompanied by unprocessed datasets, processing scripts, and readme files that specify how processing scripts need to be applied to transform unprocessed into processed data. In addition, analysis scripts will be made openly available and readme files will specify how those scripts need to be applied to reproduce the results we obtained from the published data. Where the results of TealHelix data are published in scientific journals, this documentation may also be published as supplementary material accompanying the journal publication.

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?

Each published dataset will be made available under a Creative Commons Attribution Required CC-BY 4.0 licence. For datasets that cannot be published in their entirety (e.g., because data cannot be anonymized for open sharing), TealHelix partners will make reasonable efforts to publish (under CC-BY 4.0 licence) datasets that are as rich as possible while ensuring participants' anonymity.

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

All data deposited in online repositories will be usable by third parties (with attribution required) after the end of the project. Third parties can also obtain access to unpublished datasets (provided proper ethical and legal clearance) for the duration of the retention period.

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

Each published dataset will be accompanied by documentation of data provenance and links to any reports that describe the data collection process.

Describe all relevant data quality assurance processes.

TealHelix partners will follow best practices in their respective research disciplines with regard to the collection, processing, analysis, sharing, and documentation of data. Where it is common practice in the respective partners' research field, researchers will preregister data collection activities and use established (or validated) procedures to obtain valid, reliable, and useful results that are useful to other researchers.

Each data collection activity will be coordinated within Work Package Teams. The TealHelix Coordination Team will provide input on data collection activities and approve of all deliverables.

Guidelines for the FAIR sharing of datasets will be distributed to all partners by the TealHelix Data Manager, and the Data Manager will review all datasets upon repository deposition.

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.).

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.

TealHelix research activities will generate data processing and analysis scripts, which will be deposited along with the corresponding datasets in trusted repositories. Scripts will be annotated to facilitate reuse and accompanied by readme files that specify their application. TealHelix will also involve the publication of preregistered study protocols under standard reuse licences (e.g., on the OSF). In addition, the stimulus materials, behavioural experimental procedures, and measures developed for or used in TealHelix research activities will be made openly available (in online repositories or as supplementary materials accompanying journal publications) where legally possible.

5. ALLOCATION OF RESOURCES

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)? 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).

Deposition and publication of data, scripts, preregistrations, and research materials in Zenodo and the OSF is free of charge. FAIR sharing of data and other research outputs requires non-trivial amounts of time, and this work is an essential part of the research activities of TealHelix partners.

Any costs related to the use of (secure) infrastructure for collecting, processing, storing and exchanging data during the project are covered by the host institution of the respective TealHelix partner. TealHelix includes a dedicated task to coordinate and guide the implementation of FAIR data management throughout the project.

Who will be responsible for data management in your project?

TealHelix Work Package and Task leaders are responsible for ensuring that all data are collected and managed in line with the Data Management Plan, the Project Management Plan as well as institutional and national guidelines and policies. This includes the responsibility to secure clearance from institutional review boards or ethics committees before initiating data collection.

The TealHelix Data Manager at KU Leuven (Florian Lange, florian.lange@kuleuven.be) will guide the implementation of the Data Management Plan, support project partners in establishing FAIR data sharing procedures, and review datasets before publication. The Coordination Team will provide advice in cases of doubt regarding project-related data and ethics management.

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)?

Deposited data and the accompanied metadata, documentation, and other research outputs will remain available for the lifetime of the repository. Retention periods for data that cannot be deposited publicly (e.g., personal data) will follow relevant national and institutional regulations and guidelines. Project partners will ensure that the retention of such data is in line with approval from relevant data protection authorities (institutional review boards, ethics committees). Partner institutions will cover any additional costs for long-term storage.

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 collection and management of personal data will be in line with the General Data Protection Regulation (GDPR). Clearance from institutional review boards or ethics committees will be secured to ensure that data management adheres to relevant institutional and national regulations and guidelines.

The collection of personal data will be restricted to the minimum necessary for the intended purposes, preventing unwarranted data accumulation and preserving data relevance. To collect and store personal data, secure platforms will be used that can only be accessed using multi-factor authentication. For storage, datasets will either be anonymized (i.e., through the deletion of personal data that would allow identifying individuals) or pseudonymized and encrypted/password-protected. Data will be stored on secure servers that are regularly backed up.

Before sharing datasets between institutions, data will be anonymized or the institutions will draft data transfer agreements for the exchange of personal data (according to relevant procedures at the respective institutions). Anonymized datasets and other forms of non-sensitive data will be exchanged among project partners using the TealHelix SharePoint platform at KU Leuven. This solution provides access control, backup procedures, and file versioning.

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

Anonymized datasets as well as the accompanying metadata, documentation, and other research outputs will be permanently (i.e., for the lifetime of the repository) stored in trusted repositories (Zenodo, OSF).

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).

As specified in Section 6 of the Project Management Plan (Deliverable 1.1), TealHelix partners will conduct thorough cost-benefit analyses to ensure that the societal and scientific benefits of research activities outweigh any potential risks to participants. Throughout all research activities, they will strive to maximize those benefits (e.g., through collecting robust data and facilitating data reuse) and to minimize risks (e.g., through limiting the collection of personal data and safeguarding participants' privacy). All planned research activities will be reviewed by institutional review boards or ethics committees and data collection activities will only commence after having received ethics approval.

Research participants will be informed that their data will be made publicly available in anonymized form in a trusted repository.

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

All TealHelix research activities that involve the collection of data from identifiable research participants will obtain informed consent. Consent forms will be used to inform participants that their data will be made permanently publicly available in anonymized form in a trusted repository. The precise form and content of the informed consent form will depend on the local ethics authority.

Added question: How will gender and other diversity dimensions be assessed?

Where applicable, TealHelix data collection and processing activities will take into account gender, diversity dimensions intersecting with gender, and facets of individual vulnerability. As a rule, project activities will aim at an assessment of diversity that is as comprehensive as possible while refraining from collecting unnecessary personal data. Project partners will use inclusive measures to assess sociodemographic characteristics and diversity dimensions, and respondents will not be required to respond to questions that do not apply to them. For example, respondents will be given the option to self-describe their gender identity, ethnicity, or professional status, or they can refrain from answering questions related to such characteristics.

8. OTHER ISSUES

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

TealHelix partners will take into account the Guidelines on FAIR Data Management in Horizon 2020 and the European Code of Conduct for Research Integrity to uphold ethical research practices in all research activities, as well as any institutional guidelines on research integrity, open science and data protection at the respective partner institutions.