D8.3. Preliminary Data Management Plan

Funded by the European Union

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 the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

Deliverable Information Sheet

Version	VF
Grant Agreement Number	101130073
Project Acronym	PHAntastic
Project Title	PHA-based innovative agricultural solutions to deliver biobased fertilisers and plant protection products
Project Call	HORIZON-CL4-2023-RESILIENCE-01-TWO-STAGE
Project Duration	September 2024 – August 2028 (48 months)
Deliverable Number	D8.3
Deliverable Title	Preliminary Data Management Plan
Deliverable Type	DMP
Deliverable Dissemination Level[1]	PU
Work Package	WP8
Lead Partner	IDEA
Authors	Nina Jeliazkova, Nikolay Kochev, Vedrin Jeliazkov
Contributing Partners	-
Reviewers	Dorottya Kalo (RVL), Juan Luis Rodrigues (RTDs), Carmen Fernández (CETEC)
Official Due Date	28 February 2025
Delivery Date	30 March 2025

[11] **Type** _[1] **R**=Document, report; **DEM**=Demonstrator, pilot, prototype; **DEC**=website, patent fillings, videos, etc.; **OTHER**=other_
Dissemination level [1] **PU**=Public, **CO**=Confidential, only for members of the consortium (including the Commission Services), **CI**=Classified

Revisions

Version	Date	Comments	Authors
V0	5/03/2025	First draft of document	Nina Jeliazkova, Nikolay Kochev, Vedrin Jeliazkov
V1	25/03/2025	Revised version based on the comments of the internal reviewers	Nina Jeliazkova, Nikolay Kochev, Vedrin Jeliazkov
V2	25/03/2025	Revised version based on the comments of the WP leader	Nina Jeliazkova, Nikolay Kochev, Vedrin Jeliazkov
V3	26/03/2025	Revised version based on the comments of the corresponding manager.	Nina Jeliazkova, Nikolay Kochev, Vedrin Jeliazkov
VF	28/03/2025	Final version, approved by the project coordinator and submitted to EC.	Nina Jeliazkova, Nikolay Kochev, Vedrin Jeliazkov

Document verification and approval

	Name	Organisation	Date	Visa
Internal Reviewer 1	Dorottya Kalo	RVL	20/03/2025	OK
Internal Reviewer 2	Juan Luis Rodrigues	RTDS	24/03/2025	OK
WP Leader	Carmen Fernández	CETEC	24/03/2025	OK
PHAntastic manager	Nina Jeliazkova	IDEA	26/03/2025	OK
Project Coordinator	Carmen Fernández	CETEC	27/03/2025	OK

List of Acronyms

AAM	Author Approved Manuscript
CC	Creative Commons
DOI	Digital Object Identifier
DMP	Data Management Plan
ELN	Electronic lab notebook
FAIR	Findability, Accesibility, Interoperability, Reusability
GDPR	General Data Protection Regulation
IP	Intellectual Property
IPR	Intellectual Property Rights
NDA	Non Disclosure Agreement
OA	Open Access
OECD	Organisation for Economic Co-operation and Development
REACH	Registration, Evaluation, Authorisation and restriction of Chemicals
VoR	Version of Record

List of Figures

Keywords list

- Data Management Plan
- Data collection
- Data generation
- FAIR data
- Metadata
- Data security
- Data re-use
- Ethic

Disclaimer

Table of Contents

1. Executive summary

This is the preliminary version of the Data Management Plan (DMP) of the PHAntastic project. This DMP defines the methodologies needed to ensure safe data collection, handling, and storage. IDEA will create an Intranet for research sharing, and easy-to-use data entry templates (automatically converted to FAIR machine-readable data for data analysis) and will prepare the Data Management Plan. Moreover, KVC will develop the document on Ethics principles, which will deal with ethical issues ensuring citizens' rights, security & welfare. Thus, this deliverable aims to inform the partners about the main guidelines and good practices that will be applied to the data processed and generated in the project. The initial DMP is available at DMPOnline https://dmponline.dcc.ac.uk/plans/159697 .

The DMP is intended to be a living document in which information will be available at a more detailed and specific level through regular updating as the project implementation progresses and when significant changes or milestones occur. It references Horizon Europe's recommendations for Open Science implementations.

IDEA has undertaken the creation of an electronic lab notebook to facilitate research sharing among partners. This involves developing easy-to-use data entry templates which are automatically converted into FAIR-compliant, machine-readable data formats for analysis. The templates, as shown in the projects' eNanoMapper instance, facilitate data interoperability through a step-wise approach that captures method descriptions, experimental factors, and replicates. By leveraging the FAIRification tools, ontology annotation becomes integrated, further enhancing data harmonization.

The DMP emphasizes ensuring data interoperability by employing standard vocabularies and metadata, facilitated by tools like the Template Designer. Zenodo is the preferred choice for long-term archival, providing each dataset with a DOI for persistent identification. Furthermore, to meet the needs of data reuse, eNanoMapper offers a REST API for programmatic data access and analysis. Certain datasets will remain closed due to proprietary concerns, NDAs, or commercial interests, but metadata associated with these datasets will be openly accessible under a CC0 license.

To support open science practices, the DMP encourages sharing datasets openly, under licenses like CC-BY, when possible. Training on these licenses and open science concepts is anticipated for project partners to foster data-sharing culture.
While primarily serving the project's objectives in developing safe and sustainable PHA products, the data has broader applications. Stakeholders such as regulatory bodies, manufacturers, environmental experts, and the scientific community can access the data for hazard assessments, product optimization, life cycle assessments, and bioplastic research, respectively.

Finally, the DMP ensures compliance with GDPR, keeping personal data protected. By adhering to the principles outlined in this Data Management Plan and maintaining compliance with GDPR requirements, the project aims to ensure the protection of personal data throughout its lifecycle.

2. Introduction

Managing digital research data in accordance with the FAIR principles - Findable, Accessible, Interoperable, and Reusable - is essential for ensuring data transparency, reusability, and long-term accessibility. The Horizon Europe requires researchers to deposit (meta)data as soon as possible after production, generation, or post-processing and quality control. All data should be stored in a trusted repository and made openly accessible as soon as possible. The guiding principle is "as open as possible, as closed as necessary," meaning that while data should be open by default, restrictions may apply when justified (e.g., for ethical, legal, or confidentiality reasons).

At the moment of writing 5 trusted repositories are 'ready' for compliance (HAL, AUSSDA, <intR>²Dok, DANS Data Station Archaeology, Zenodo). Many of the existing repositories are currently making changes and status can be checked at https://zenodo.org/records/13919643. The recommendation is "if in doubt, use Zenodo". Research data should ideally be shared under an open license, such as CC-BY or CC0, to maximize reuse and facilitate proper attribution. Additionally, to ensure the data's reusability and allow others to validate or replicate the research, it is essential to provide detailed information about the research outputs, including tools or instruments necessary for reusing or validating the data. This can include data, software, algorithms, protocols, models, workflows, and even electronic notebooks.

Valid justification for not opening the data might include (among others):

- Commercially valuable data if it would undermine its exploitation or other results (e.g. endanger trade secrets ('soft' IP)), or make IP protection of results more difficult.
- Data protection/privacy rules of sensitive and/or personal data.
- Security rules for projects dealing with strategic assets, interests, autonomy or security of the EU.

3. Data Management Plan

The initial DMP is available at DMPOnline https://dmponline.dcc.ac.uk/plans/159697

3.1. Data Summary

The summary addresses the following issues:

3.1.1. Purpose of the data collection/generation

The data collection is aimed at several key objectives, including conducting a life cycle analysis of PHAntastic products, characterizing active bioproducts for mulch films and foams, collecting information for the Safe and Sustainable by Design (SSbD) study, and gathering data from chemical and biological analyses. Additionally, data is generated for formulation and compounding of PHA blends, assessing material formation and properties, obtaining scientific data for research development, evaluating social sustainability and end-user acceptance, ensuring process suitability and replicability, and validating and testing developed products in field tests.

3.1.2. Relation to the objectives of the project

The collected data supports the project objectives by evaluating the environmental impact of PHAntastic products, characterizing essential active bioproducts, and ensuring the safety and sustainability of the proposed innovation. It enables the validation of products in terms of their end-of-life and environmental safety, ensures the production of PHA blends with required properties, and aids in generating capsules and foams. The scientific data is crucial for achieving research goals, assessing social sustainability, ensuring process feasibility, and validating product efficacy through field testing. Ultimately, the data collection contributes to delivering innovative, sustainable, and effective PHA-based delivery systems.

3.1.3. Types, formats and size of data generated/collected

The project generates various types of data, including raw data, processed data, numeric results, spectra, images, and reports. Additionally, questionnaires, interviews, and focus groups are used for data collection when necessary. Partners store data in multiple formats, including Excel spreadsheets, PDFs, text files, images, and specialized software formats for analytical data. Some partners collected data through internal datasheets, and two partners use internal Laboratory Information Management Systems. The expected data sizes range from kilobytes (kB) to gigabytes (GB), depending on the data type and format.

3.1.4. Re-using existing data

Existing data will be reused in specific areas of the project. Data from chemical and material databases, REACH dossiers, and other sources will be utilized for material selection and safety assessment. Additionally, benchmark data from existing databases will support various analyses, including Life Cycle Assessment inventories. Previous internal research data will also be referenced, and some existing data will aid in scaling up PHA blends by August 2028.

The inventory of existing data will be updated accordingly during the course of the project.

3.1.5. Outline the data utility: to whom will it be useful

The data is primarily intended for project purposes, supporting the development, validation, and assessment of PHA-based products (safe-and-sustainable-by-design).

However, the data may also be useful to other stakeholders, including regulatory bodies (e.g., for hazard assessment and REACH compliance), industry and manufacturers (for scaling up and optimizing formulations), environmental experts (for Life Cycle Assessments), end users (for product acceptance and safety), and the broader scientific community (for research in bioplastics and sustainability).

3.2. FAIR data

3.2.1. Making data findable, including provisions for metadata

This principle prescribes that the research output must be findable. That is, effort must be made to make sure people can find the data. To be findable, the principles specify:

- *F1. (meta)data are assigned a globally unique and persistent identifier*
- *F2. data are described with rich metadata*
- *F3. metadata clearly and explicitly include the identifier of the data it describes*
- *F4. (meta)data are registered or indexed in a searchable resource*

After data collection on the projects shared folder, data will be indexed and stored in PHANTASTIC eNanoMapper database (part of https://enanomapper.adma.ai [1]), with dataset archival on Zenodo, which is where dataset DOI will be assigned.

3.2.1.1. Outline the discoverability of data (metadata provision)

Partner responses indicate varying levels of data discoverability importance, with most ratings between 3 and 4 (on a scale where 4 suggests high importance). While non-confidential information will be primarily be available through the PHANTASTIC database (***Figure 1***) and the PHANTASTIC project website. Additionally, some data may be accessible through external platforms like PubChem, PubMed, journal publications, and Google Dataset Search. However, the discoverability of specific datasets will depend on their confidentiality and publication policies. Typical search queries will likely include material types, material identifiers, assay/experiment types, industrial biotechnology, PHBV production, and biopolymer production. Some data will be accessible via company websites, where information on raw materials, PHA blends, and formulation services may be provided.

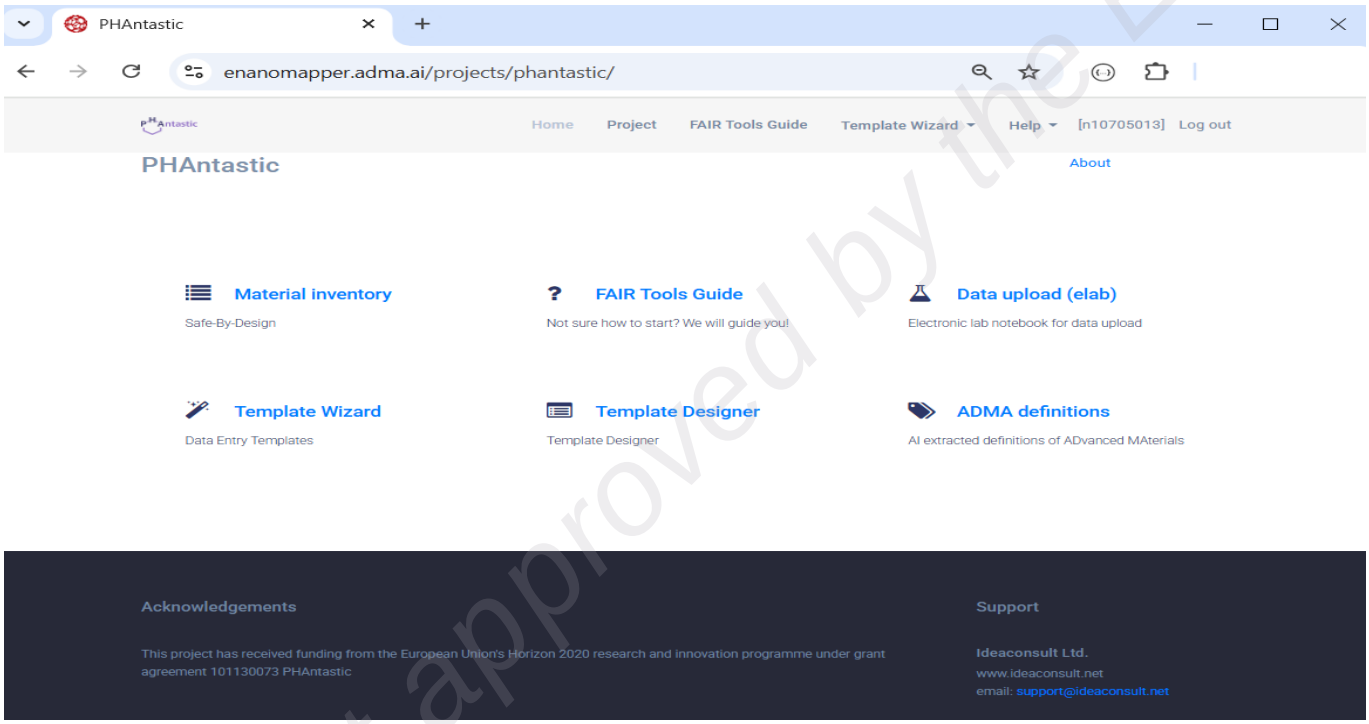

Figure 1. *Start page of the PHANTASTIC database at https://enanomapper.adma.ai/projects/phantastic/*

3.2.1.2. Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

Yes, we will consider European Registry of Materials identifiers [2] and other existing material identifiers (e.g. registry identifiers like CAS and EC numbers where available) and machine readable systems for identifying chemical compounds (e.g. SMILES, InChI). -Datasets archived on Zenodo repository will be assigned a DOI.

3.2.1.3. Outline naming conventions used

We will follow the existing eNanoMapper database FAIRification workflow [3] and ontology [4] to annotate characterisation methods and process parameters with ontology identifiers.

3.2.1.4. Outline the approach towards search keyword

eNanoMapper provides free text and faceted search at https://enanomapper.adma.ai , which is capable of aggregating search results from local and remote databases (based on the widely used Apache Solr [2]search engine). We will extend the existing search facilities with enabling domain specific search (e.g. material and spectra search) and introducing modern AI approaches allowing to combine text search and structured data queries. Metadata of all eNanoMapper databases is available at CC0 license and indexed at Google Dataset Search [3].

3.2.1.5. Outline the approach for clear versioning

Each data entry in eNanoMapper database(s) is assigned with a release version tag

3.2.1.6. Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

Data will be generated and collected according to standards relevant to the activity and in line with OECD and ISO recommendations. The eNanoMapper ontology[4] (and extensions) are used to annotate e.g. material types, endpoints, methods and other relevant parameters.

The PHANTASTIC eNanoMapper database has an integrated Template Wizard tool [5], which provides customizable data entry templates , already harmonized between several projects and with integrated ontology annotation and material identifiers.

[2] https://solr.apache.org/

[3] https://datasetsearch.research.google.com/

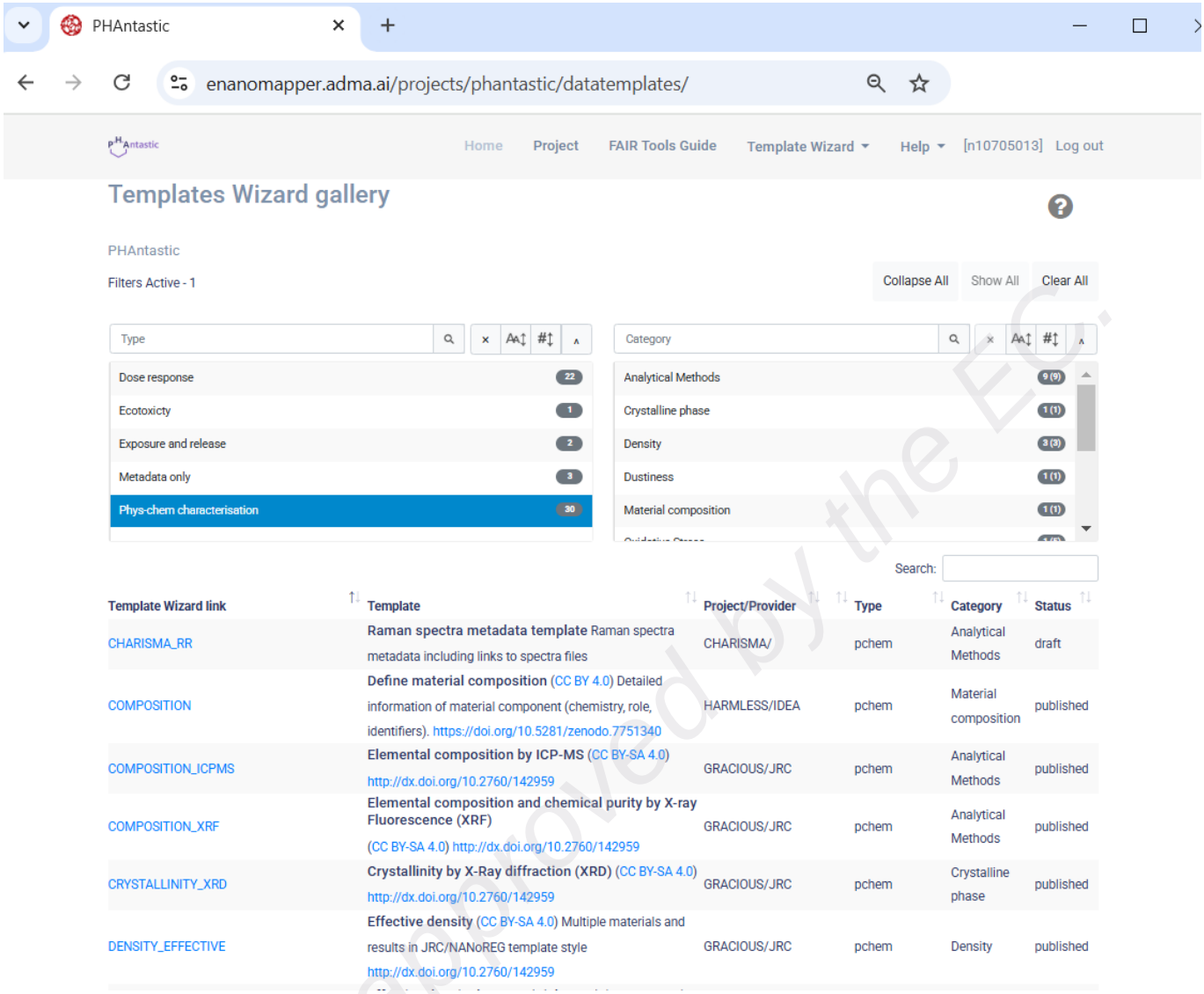

Figure 2. Template Wizard at https://enanomapper.adma.ai/projects/phantastic/datatemplates/

The Template Designer is a new online tool (***Figure 3***) allows to design templates for arbitrary types of experiment. The templates can be downloaded as Excel files, filled with by partners, saved in a shared folder and subsequently converted to machine readable form and indexed by eNanoMapper database and/or uploaded to trusted repository, as required by HORIZON Europe.

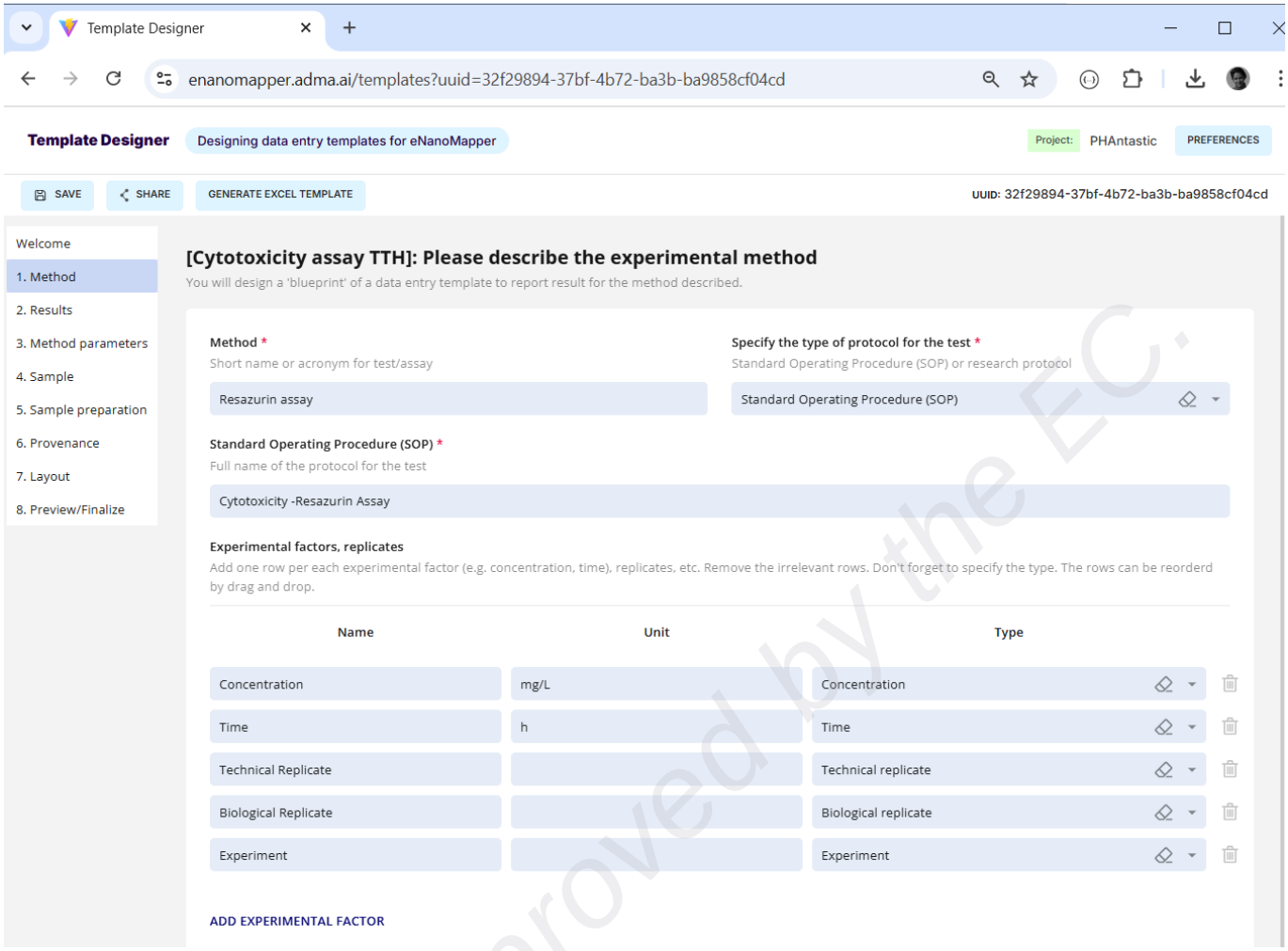

Figure 3. Template Designer at https://enanomapper.adma.ai/templates/ . Example with a dose-response assay provided for illustration.

The Template Designer starts with a Welcome screen. Besides entering template name and acknowledgement (e.g. to a project, organisation, etc) and author name, one has to specify if he/she is a lab researcher and/or data manager. Some of the questions are mandatory or optional depending on these roles. Any user can design a draft template blueprint, but only a data manager can finalize one. The Draft and finalized template blueprints appear at separate tabs at the front page. The finalized template blueprints cannot be changed anymore, but can be copied into new draft blueprints. All templates can be shared through web page links and of course all can be used to generate Excel templates.

The next step is to describe the method (***Figure 3***). The steps are accessible at the left, but one can't jump into subsequent steps unless the mandatory information is filled in. The Method page lets the user to enter the method name, specify the type of the SOP, and in case of research protocol, provide details of the research protocol, including links to external documents.

An important information to be entered in the Method page is to specify the experimental factors and the presence of replicates – these will be used when generating the result data sheets. By default, the concentration, time and replicates are filled in, but the table is editable and rows can be removed, added or modified. At the bottom of the page there are two linked dropdown boxes, allowing to specify the type of the experiments, as in eNanoMapper database. These categories are closely following the OECD Harmonized Templates classification, and where relevant are extended with ontology entries.

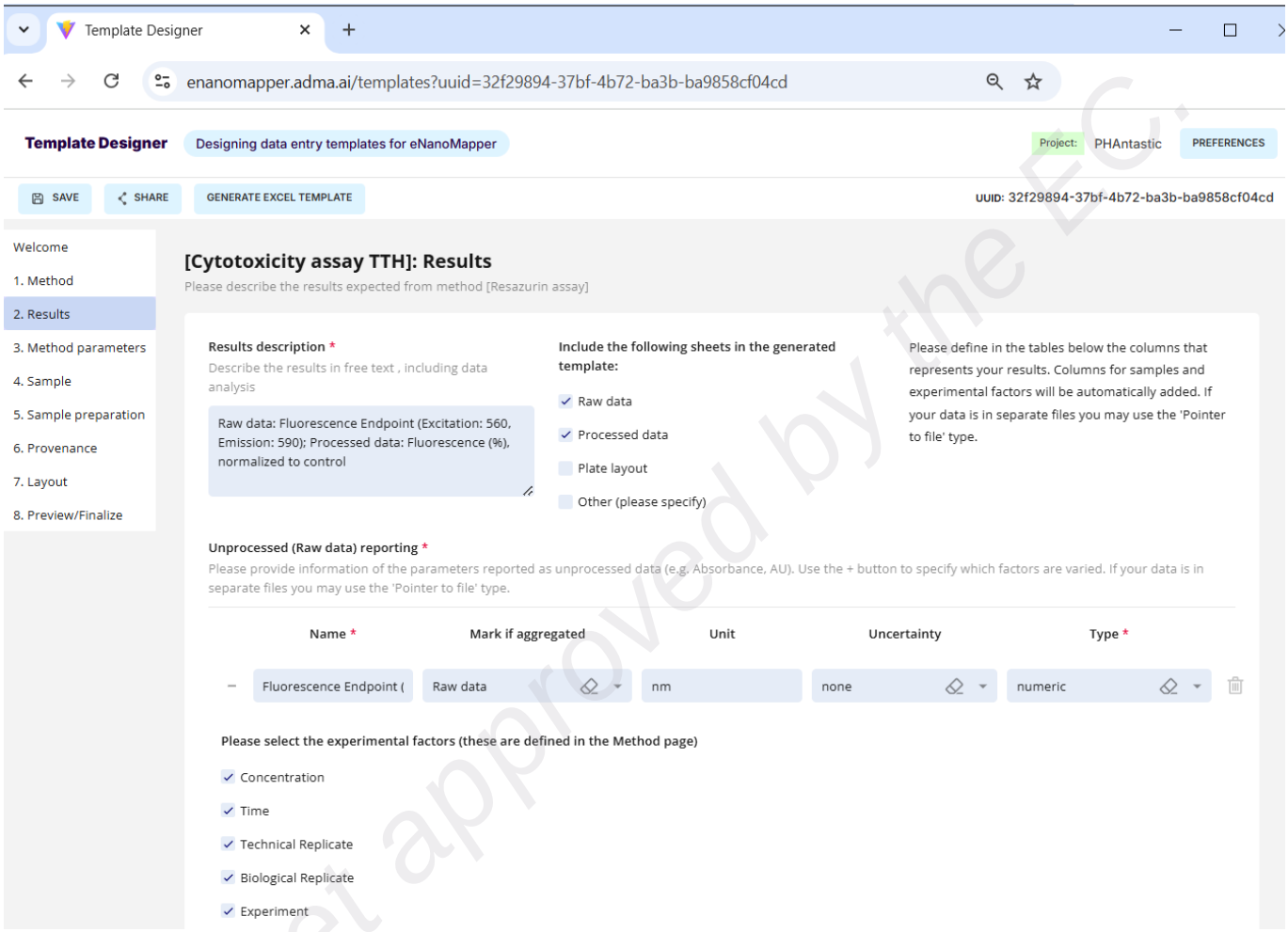

Figure 4. Template Designer – Results page. The table defines what results will be reported as raw (top table) and processed data (not visible).

The results page (*Figure 4*) allows to define the type and content of the results reported. The table defines what results will be reported as raw and processed data. For each row one can define which experimental factors will be varied.

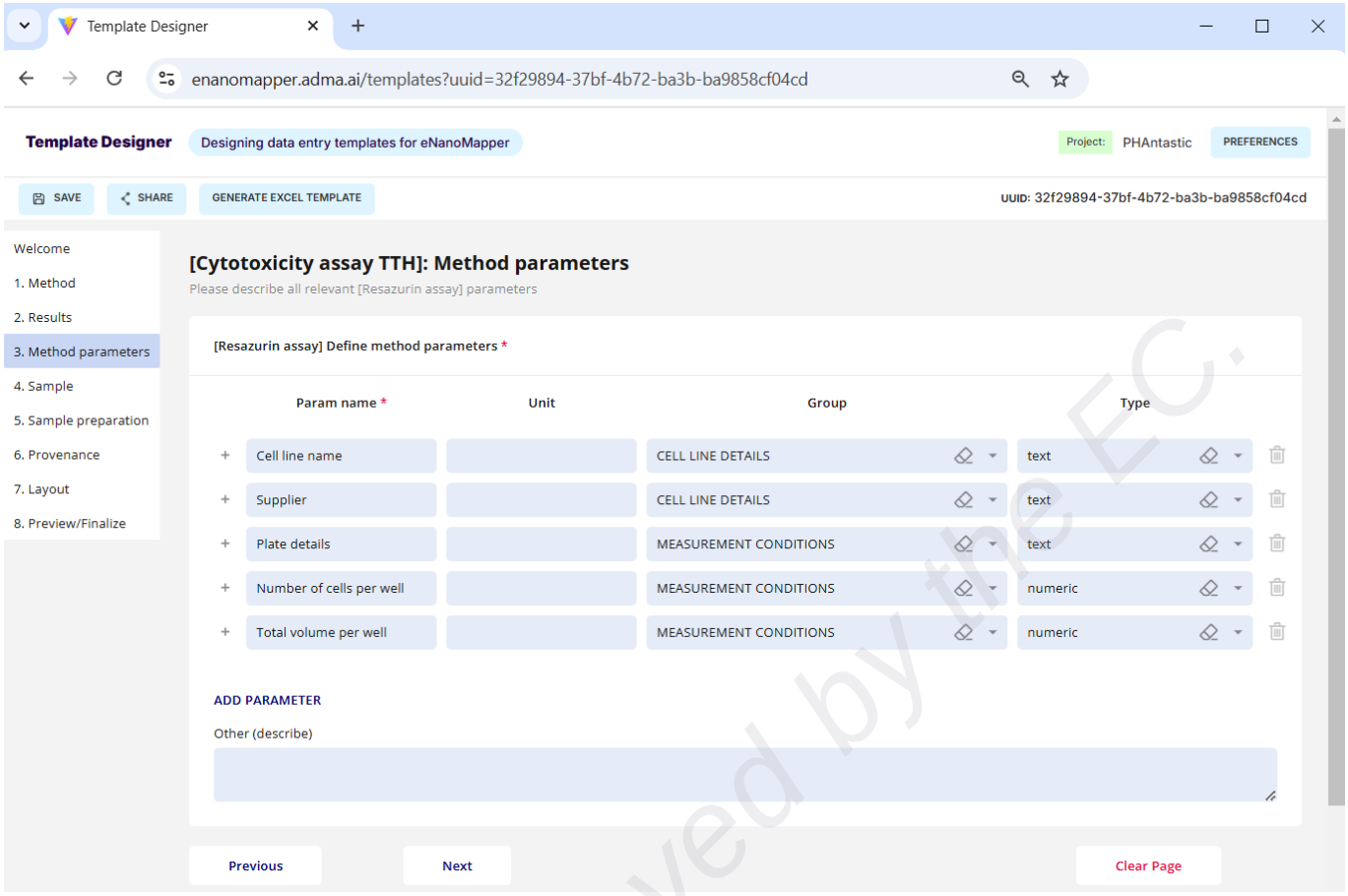

Figure 5. Template Designer – Methods parameters.

Generic metadata for the experiment (who and when did the experiment, Section 6). Provenance is provided by fixed fields; some of them can be configured to have default values. These can also be changed in the generated Template. Finally, the preview page asks for the author ORCID. This is mandatory only to finalize the draft.

On each stage there is a button “Generate Excel template” which will download the template according to the design.

3.2.2. Making data openly accessible:

This principle prescribes that data must be accompanied with metadata that explains how people get access to the data. This does not imply they always get access (generally they do), but if they do, how. To be accessible, the principles specify:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1 the protocol is open, free, and universally implementable
- A1.2 the protocol allows for an authentication and authorization procedure, where necessary

- A2. metadata are accessible, even when the data are no longer available.

3.2.2.1. Specify which data will be made openly available?

Based on partners responses, most data generated within the PHANTASTIC project will be used internally and may not be openly available due to proprietary concerns, NDAs, or commercial interests. Some partners explicitly stated that data reuse should be limited, and certain datasets will only be accessible within the project consortium.

3.2.2.2. Specify where the data and associated metadata, documentation and code are deposited

Horizon Europe requires data to be shared in trusted repositories (currently there are limited list of such repositories, e.g. Zenodo), with preferably CC-BY license (non-commercial licenses are not favoured). Data may stay restricted (because of IPR reasons), but still need to be FAIR and metadata is required to be shared openly under CC0 license (i.e. public domain). Personal data is not to be shared.

The planned open-source code for polymer property prediction will be hosted at GitHub https://github.com/ideaconsult

Based on partner responses the familiarity with Creative Commons licenses is low, therefore we plan to provide training for FAIR concepts and the relevant Horizon Europe requirements.

3.2.2.3. Specify how the data will be made available

Partners have own procedures and storage space for the data generated. File formats differ, from Excel files to file formats specific to instruments and commercial software being used. To ensure interoperability, we will propose tools to convert to harmonized data formats, annotated with appropriate metadata. The data will be indexed through the PHANTASTIC database and dataset in machine readable formats archived in Zenodo.

The PHANTASTIC eNanoMapper database will serve as the central indexing and storage platform, with access restricted to project partners. The eNanoMapper role-based access will be applied if a dataset has to be shared only within limited group of project partners and not the entire consortium.

An internal online form will be provided to declare datasets that are planned to be shared, similar to the process for publications, ensuring systematic tracking and management of project data. Therefore, data sharing on public repositories, such as Zenodo, will be conducted based on consortium agreements regarding licensing and embargo periods. Open-access data will follow FAIR (Findable, Accessible, Interoperable, Reusable) principles and be deposited in recognized repositories, including Zenodo and relevant domain-specific platforms.

3.2.2.4. Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open-source code)?

The PHANTASTIC eNanoMapper instance is hosted by IDEA and is accessible through secure HTTP connection at https://enanomapper.adma.ai/projects/phantastic/. It will include data and metadata. The data can be accessed via web browser or programmatically (REST API at https://api.ideaconsult.net) that can be used for data analysis. There is an opensource library to access eNanoMapper database and examples.

3.2.2.5. Specify how access will be provided in case there are any restrictions

PHANTASTIC eNanoMapper database access is restricted to project partners in order to support implementation of project tasks. –PHANTASTIC partners will decide which dataset and when can be shared, will specify the appropriate license and subsequently archived at a trusted repository (Zenodo).

3.2.3. Making data interoperable:

This principle prescribes that data must be sufficiently annotated such that people can understand what the data means. This means a clear data format must be used, values must have units, columns must have clear annotation what was measured, etc. This is also where ontologies (vocabularies) typically come in. To be interoperable, the principles specify:

- *I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*
- *I2. (meta)data use vocabularies that follow FAIR principles*
- *I3. (meta)data include qualified references to other (meta)data*

3.2.3.1. Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

Proper data exchange formats will be selected, based on the existing practice and solutions listed in previous sections. It is essential to keep the data life cycle in mind here. For short term data exchange, the tools should be convenient to use in daily lab work, therefore suggesting the template tools explained in previous sections. To get data indexed in a search engine more application-oriented format may be more appropriate, therefore the templates should be machine readable and not arbitrary Excel files. For long term data archival a format is preferred that is self-contained and self-explanatory, for this purpose we will offer conversion to NeXus format https://www.nexusformat.org/ , which is one of the recommended formats in multiple domains. If required, export into OECD Harmonized Templates can be considered. These formats do not exclude each other, but are rather complementary.

3.2.3.2. Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

The use of a vocabulary is an important aspect of data harmonisation. According to survey responses, ontologies are not used on daily basis, therefore training and user-friendly tools to adoption of ontology annotation will be required. The eNanoMapper FAIRification tools offer assistance of ontology annotation. Upon conversion to machine readable format the terms are annotated with ontologies. The eNanoMapper ontology has a well-defined procedure for adding terms from existing ontologies [4], which will be followed when relevant terms and identified (e.g. from EPPO [6])

3.2.4. Increase data re-use (through clarifying licenses)

This principle outlines that community practices for data sharing and data science should be followed. To be reusable, the principle specify:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage license
- R1.2. (meta)data are associated with detailed provenance
- R1.3. (meta)data meet domain-relevant community standards

Horizon Europe requires data to be shared in trusted repositories (currently there are limited list of such repositories, e.g. Zenodo), with preferably CC-BY license (non-commercial licenses are not favoured). Horizon Europe mandates immediate open access of publications and does not allow statements “data is available by contacting the author”.

Data may stay restricted (because of IPR reasons, as defined in Annex 5 of the Grant Agreement), but still need to be FAIR and metadata is required to be shared openly under CC0 license (i.e. public domain). Personal data is not to be shared. Based on partner responses the familiarity with Creative Commons licenses is low, therefore we plan to provide training for FAIR concepts and the relevant Horizon Europe requirements.

According to partners survey responses:

- There is potential for the data to be used in scientific publications, which would make certain data available to the wider scientific community.

[4] https://enanomapper.github.io/tutorials/Added%20ontology%20terms/README.html

- The re-use of some data depends on the exploitation plan, and whether external demand for the data arises or if it is allowed.

3.3. Allocation of resources

Efficient data management and open science practices require active involvement from all partners. WP8 includes dedicated effort allocations to ensure these practices are embedded in daily research activities:

CETEC (33 PMs) - Coordinates and monitors overall project progress, ensuring data management integration aligns with the work plan. IDEA (6 PMs) leads Task 8.3 on data and ethics management, developing the intranet for research sharing, data entry templates, and ensuring compliance with FAIR principles. IDEA also prepares the Data Management Plan and supports partners in structured data handling. KVC (6 PMs) develops the ethics framework, ensuring data security, privacy, and compliance with citizens' rights. Other partners (1-3 PMs each) are expected to actively contribute by integrating data management into their daily activities. By promoting consistent use, the ELN for documentation of the SSbD workflows and internal sharing, we envisage that data management becomes a routine part of daily research activities.

3.4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

Google drive maintained by coordinator (CETEC) with the purpose of supporting project management. Only invited members can upload, download and view the content.

-FAIR database for data sharing of experiments, materials and all information related to SSbD process (IDEA). To support structured data uploads, an electronic lab notebook (ELN) has been installed at https://elab.phantastic.adma.ai . The ELN enables systematic documentation and management of experimental data while ensuring traceability and compliance with project data policies. Access is granted to project partners only, using same login credentials for both database and the ELN Technically, the data access is restricted via role based security and standard authentication and authorization protocols (API key, OAuth2), based on the open source Keycloak identity manager https://www.keycloak.org/ installed at https://iam.ideaconsult.net. By default, there is project-wide role “phantastic”, if necessary additional roles can be introduced in order to restrict access to specific group of partners.

When logging in through web browser, users are presented with a login form (**Figure *6***). The user can also use an API key for programmatic access (e.g. for data retrieval and analysis with).

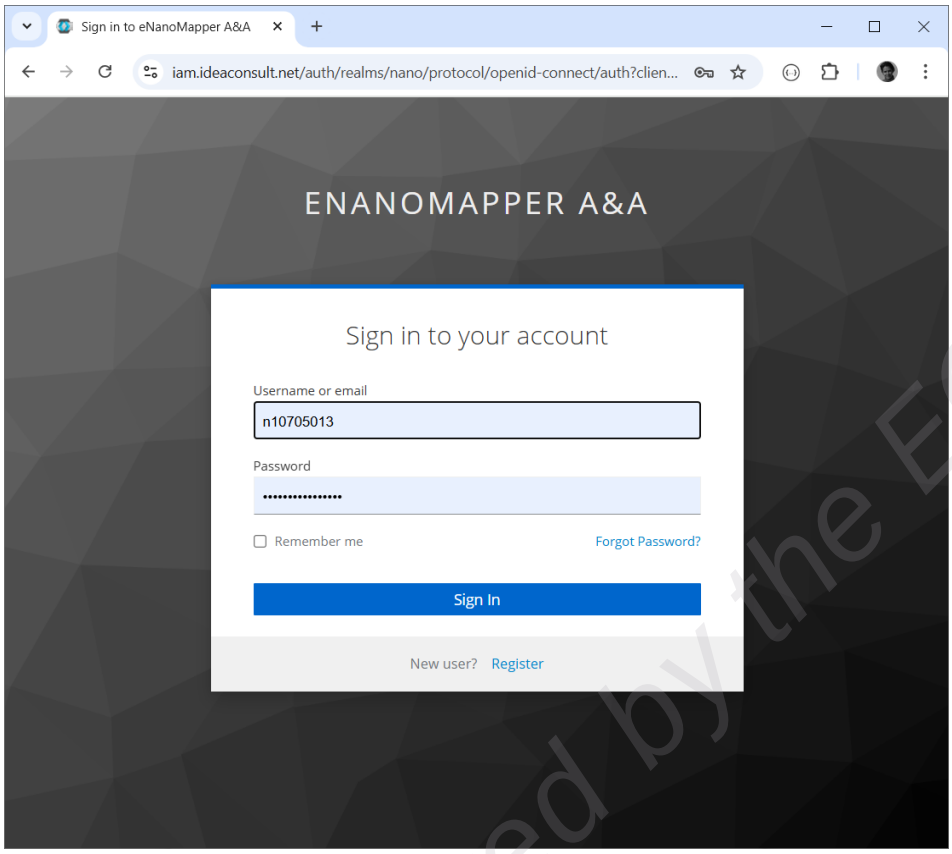

Figure 6. eNanoMapper logging form

3.5. Ethical aspects

The ethics principles outlined in the PHAntastic deliverable D8.1 are grounded in a comprehensive framework aimed at ensuring ethical integrity throughout the project. These principles are informed by the Belmont Report's ethical guidelines, emphasizing respect, beneficence, and justice:

- Respect for Persons: This principle underscores the importance of treating individuals as autonomous agents and includes measures to protect the autonomy and rights of vulnerable individuals.

- Beneficence: Ensures that participants are protected from harm while maximizing possible benefits and minimizing potential harms, particularly for those in situations of material deprivation or social disadvantage.

- Justice: Promotes fairness in the distribution of benefits and risks of research, ensuring equitable treatment for all participants.

The key actions taken to ensure adherence to ethical principles are summarised below:

- Ethical Guidelines and Compliance: The project establishes comprehensive guidelines for all partners to address ethical issues within the project framework, ensuring compliance with ethical standards and relevant legislation. This includes informed consent, research integrity, and data protection.

- Informed Consent and Participant Protection: Informed consent is a critical component to respect participants' autonomy. The project requires explicit, freely given consent from participants regarding any data collection and involvement in project activities, ensuring transparency and the right to withdraw without any repercussions.

- Stakeholder Engagement and Responsible Research and Innovation (RRI): The project prioritizes stakeholder engagement, particularly through participatory involvement and public engagement activities. Strategies include community readiness assessments, workshops, and citizen science projects, fostering ownership of project results among diverse groups.

- Social Sustainability and Gender Integration: Through Work Package 5 (WP5), social sustainability assessments are conducted, focusing on the impacts of project solutions on communities. Gender analysis is integrated into all project phases to ensure women's perspectives and contributions are considered, aiming to improve inclusivity and participation in science.

- Regulatory Compliance and Risk Assessment: Ensuring compliance with various EU and international regulations related to environmental risks, waste management, fertilizers, and pesticides is integral. Tasks like T5.1 prioritize regulatory compliance, assessing risks to human health and safety related to the production processes.

- Open Science and Collaboration: The project fosters open science practices to promote transparency, sharing objectives, methodologies, and results with appropriate stakeholders while protecting intellectual property rights.

- Continuous Ethical Assessment: Regular evaluations of ethical risks are performed, integrating tools and strategies to effectively translate ethical principles into actionable practices. This approach ensures the ongoing relevance and adaptation of ethical considerations throughout the project lifecycle.

Data privacy concerns across all engagement activities, including data collection from websites, surveys, and community initiatives, will be addressed through collaboration with WP7 and WP4, while technical grounds of data security stay the same.

D8.1 also includes a risk-benefit analysis to ensure that its ethical principles are effectively translated into actionable practices. Key risks identified include unforeseen ethical issues, suboptimal stakeholder engagement, regulatory hurdles, potential adverse impacts, and inadequate social sustainability assessments. The project addresses these risks with strategies such as utilizing SSH experts, implementing stakeholder plans, ensuring regulatory compliance, and conducting thorough sustainability assessments. The anticipated benefits of the PHAntastic project include promoting environmental sustainability by reducing reliance on traditional agricultural plastics and agrochemicals, enhancing biodegradability, and fostering interdisciplinary collaboration. Additionally, the project aims to involve end-user perspectives, particularly horticultural and fruit producers, to ensure that the bioproducts developed are both socially and economically beneficial. This comprehensive assessment highlights PHAntastic's commitment to ethical research by prioritizing human health, environmental conservation, and societal engagement while maintaining transparency and accountability in its processes.

3.6. Other

Horizon Europe provided recommendations for Open Science requirements implementations at https://rea.ec.europa.eu/open-science_en . The example DMP template provided is similar to this one.

We are aware of DMP tool promoted by the funder and may switch subsequent versions of this plan to https://argos.openaire.e u which will provide integration with Zenodo and the new EOSC Node (launched 11 Oct 2024) .

IPCHEM (https://ipchem.jrc.ec.europa.eu/ , an EC promoted database for biomonitoring) and European Platform on LCA https://eplca.jrc.ec.europa.eu/ were mentioned in the proposal and will be considered for the relevant data.

4. Conclusions

An initial Data Management plan is described, addressing the FAIR criterial and structured questions from DMP Online. Specific tools and solutions are envisaged to enable data management as part of daily collaborative work to implement SSbD.

The DMP defines methodologies for safe data collection, handling, and storage, aligned with Horizon Europe's open-access mandates. By leveraging tools like the eNanoMapper database, the plan facilitates the transformation of data into FAIR-compliant, machine-readable formats, thereby enhancing data utility beyond the project. The provision of metadata ensures that even when data access is restricted due to proprietary concerns or NDAs, the essential information remains discoverable under a CC0 license.

The key challenges identified include managing proprietary data sensitivity, facilitating partner involvement and data interoperability, and ensuring ethical compliance in data handling. The project has addresses these through several initiatives:

- To balance openness with data sensitivity, The plan involves using the eNanoMapper tools (electronic lab notebook https://elab.phantastic.adma.ai/ and https://enanomapper.adma.ai/projects/phantastic/) for internal access, with selected datasets becoming available post-IPR clearance and archived in trusted repositories like Zenodo.

- Acknowledging diverse data formats among partners, the DMP proposes using existing conversion tools to harmonize data into machine-readable formats. The adoption of standard vocabularies and ontology annotation through FAIRification tools further facilitates cross-disciplinary data exchange and reuse.

- By integrating ethical considerations into data management processes, the DMP complies with GDPR, ensuring personal data protection. The implementation of informed consent processes further underscores the commitment to ethical research practices.

The DMP supports PHAntastic's core project objectives, namely developing safe and sustainable PHA-based solutions. The collected data is essential for tasks involving the material characterization, hazard assessment, life cycle analysis, and sustainability assessments. By ensuring that data is effectively managed and utilized, the DMP contributes to reinforcing the project's scientific rigor and fostering interdisciplinary collaboration.

By promoting data transparency and accessibility, the DMP aids regulatory compliance tasks, supports enhanced stakeholder engagement. The plan is aligned Horizon Europe's requirements for open science.

5. References

[1] N. Jeliazkova, M. D. Apostolova, C. Andreoli, F. Barone, A. Barrick, C. Battistelli, C. Bossa, A. Botea-Petcu, A. Châtel, I. De Angelis, M. Dusinska, N. El Yamani, D. Gheorghe, A. Giusti, P. Gómez-Fernández, R. Grafström, M. Gromelski, N. R. Jacobsen, V. Jeliazkov, K. A. Jensen, N. Kochev, P. Kohonen, N. Manier, E. Mariussen, A. Mech, J. M. Navas, V. Paskaleva, A. Precupas, T. Puzyn, K. Rasmussen, P. Ritchie, I. R. Llopis, E. Rundén-Pran, R. Sandu, N. Shandilya, S. Tanasescu, A. Haase, P. Nymark, *Nat. Nanotechnol.* **2021**, *16*, 644–654.

[2] J. van Rijn, A. Afantitis, M. Culha, M. Dusinska, T. E. Exner, N. Jeliazkova, E. M. Longhin, I. Lynch, G. Melagraki, P. Nymark, A. G. Papadiamantis, D. A. Winkler, H. Yilmaz, E. Willighagen, *J. Cheminform.* **2022**, *14*, 57.

[3] N. Kochev, N. Jeliazkova, V. Paskaleva, G. Tancheva, L. Iliev, P. Ritchie, V. Jeliazkov, *Nanomaterials* **2020**, *10*, 1908.

[4] J. Hastings, N. Jeliazkova, G. Owen, G. Tsiliki, C. R. Munteanu, C. Steinbeck, E. Willighagen, *J. Biomed. Semantics* **2015**, *6*, 10.

[5] N. Jeliazkova, E. Longhin, N. El Yamani, E. Rundén-Pran, E. Moschini, T. Serchi, I. V. Vrček, M. J. Burgum, S. H. Doak, M. R. Cimpan, I. Rios-Mondragon, E. Cimpan, C. L. Battistelli, C. Bossa, R. Tsekovska, D. Drobne, S. Novak, N. Repar, A. Ammar, P. Nymark, V. Di Battista, A. Sosnowska, T. Puzyn, N. Kochev, L. Iliev, V. Jeliazkov, K. Reilly, I. Lynch, M. Bakker, C. Delpivo, A. Sánchez Jiménez, A. S. Fonseca, N. Manier, M. L. Fernandez-Cruz, S. Rashid, E. Willighagen, M. D Apostolova, M. Dusinska, *Nat. Protoc.* **2024**, DOI 10.1038/s41596-024-00993-1.

[6] A. Ayllón-Benitez, J. A. Bernabé-Diaz, P. Espinoza-Arias, I. Esnaola-Gonzalez, D. S. A. Beeckman, B. McCaig, K. Hanzlik, T. Cools, C. Castro Iragorri, N. Palacios, *Front. Artif. Intell.* **2023**, *6*, DOI 10.3389/frai.2023.1131667.