Disclaimer:

This deliverable has been submitted to the European Commission and is currently under review. The final version after the approval may differ.

D1.1 - Data Management Plan

Name of the author   Karolína Řípová, Pedro Augusto
Organization   GRANT Garant, CENTRALESUPELEC
Date of the presentation   29.04.2024

Co-funded by
the European Union

D1.1 – Data Management Plan (DMP)

Project Acronym:	**PROTEIN4IMPACT**
Project Title:	**IMPACT OF ALTERNATIVE PROTEIN SOURCES TO IMPROVE NUTRITION**
Grant Agreement:	**4101182324**
Project coordinator:	**CENTRALESUPELEC (CS), PIC 933906883, PLATEAU DE MOULON 3 RUEJOLIOT CURIE, GIF SUR YVETTE 91192, France**
Call:	HORIZON-CL6-2024-FARM2FORK-01
Topic:	HORIZON-CL6-2024-FARM2FORK-01-07
Type of action:	HORIZON Research and Innovation Actions
Granting Authority:	European Research Executive Agency

Deliverable D1.1 - Data Management Plan (DMP)

Short summary:	This deliverable is the first version of the Data Management Plan (DMP) for the PROTEIN4IMPACT project, funded under Horizon Europe. It outlines the management of project resulting datasets and deliverables following FAIR principles and EU Open Science policy. The DMP ensures data findability, accessibility, interoperability, and reusability, using trusted repositories (e.g. Zenodo). Sensitive datasets are protected for commercial exploitation, while open-access outputs are licensed under Creative Commons. This document serves as a living platform for continuous updates throughout the project's lifecycle.

Due date:	30/04/2025
WP, leader:	WP1, CS
Authors:	WP1, GG – Karolína Řípová
	WP1, CS – Pedro Augusto

Dissemination Level

PU	Public	☒
SEN	Sensitive	☐

Document History

Version	Date	Approved	Chapters edited	Reason for change
V1	19/03/25	Karolína Řípová	All	Document prepared
V2	09/04/25	Karolína Řípová	All	Update based on data collection among partners

List of participants

Beneficiary name	Short name
CENTRALESUPELEC	CS
UNIVERSITAET HOHENHEIM	UOH
VENUSROUZES LABSOLUTIONS	VRLS
PANEPISTIMIO DYTIKIS ATTIKIS	UNIWA
ETHNICON METSOVION POLYTECHNION	NTUA
UNIVERSIDADE DO MINHO	UMINHO
DANMARKS TEKNISKE UNIVERSITET	DTU
AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE	ENEA
FUNDACION CARTIF	CARTIF
GOLEM – GESELLSCHAFT FUR INTEGRIERTE MIKROELEKTRONISCHE KOMPLETTLOESUNGEN GMBH	GOLEM
AQUABIOTECH LIMITED	AQB
NORDIC DIASPORA FORUM	NDF
KYPRIAKOS SYNDESMOS KATANALOTON	CCA
GRANT GARANT SRO	GG
CONSIGLIO NAZIONALE DELLE RICERCHE	CNR
AGRICLIMA CONSULTING AB	AqC
IDENER RESEARCH & DEVELOPMENT AGRUPACION DE INTERES ECONOMICO	IDENER
CONSORZIO PER L'INNOVAZIONE E LA BIOECONOMIA	In.Bio

TABLE OF CONTENTS

LIST OF FIGURES

1   Summary

This document presents the data management strategy for the PROTEIN4IMPACT project, a pioneering initiative supported under the Horizon Europe programme's call HORIZON-CL6-2024-FARM2FORK-01-07. The project is coordinated by CentraleSupelec (CS). The PROTEIN4IMPACT data management plan (DMP), developed by the task leader GRANT Garant s.r.o. (GG), serves as a key strategic deliverable with public dissemination status. It acts as a blueprint for outreach, and utilization of research data produced throughout the project's lifecycle. As a living document, the DMP will be regularly revised and refined to accommodate the project's developments and necessary modifications. It establishes a structured approach to assist the consortium in promoting research and innovation efforts, ensuring compliance with Article 17 and Annex 5 of the project's Grant Agreement. Furthermore, it defines essential guidelines and responsibilities, functioning as a reference manual for management of the research data for members of the project consortium.

The PROTEIN4IMPACT project is dedicated to assessing the nutritional value, health implications, safety, and quality of novel protein-based foods (NFs) derived from unconventional sources while also evaluating their environmental and socio-economic effects. To ensure the project's results effectively utilized and achieve maximum impact, a comprehensive strategy for the management of resulting project data has been established:

PROTEIN4IMPACT Data Management Plan (DMP)

The PROTEIN4IMPACT Data Management Plan (DMP), a key output of Work Package 1 (WP1) prepared by the task leader GRANT Garant (GG), ensures the responsible handling of all research data and results in alignment with the FAIR principles (Findable, Accessible, Interoperable, Reusable) and Open Access (OA) guidelines. Key elements of this deliverable include:

Definition of Generated Datasets: The project will produce a diverse range of datasets and research outputs (public deliverables) across its seven work packages.

Findability: Public datasets will be primarily stored in Zenodo (within the PROTEIN4IMPACT community: https://zenodo.org/communities/PROTEIN4IMPACT), utilizing persistent identifiers (such as DOIs) and Creative Commons licenses (CC BY for data, CC0 for metadata). Confidential datasets will be managed under strict data protection protocols.

Accessibility: Most of the project's datasets and research outcomes will be made openly available. However, certain sensitive deliverables will remain restricted to safeguard intellectual property for potential patent applications.

Interoperability: Data will be provided in widely used open formats (such as MS Office, CSV, FASTA, and PDF) to ensure compatibility with common tools. Proprietary formats will be supported by licensed software or Python libraries where necessary.

Reusability: Comprehensive metadata, README files, and descriptive documentation will accompany datasets, ensuring ease of access, reproducibility, and compliance with Horizon Europe requirements.

Updates: The DMP, prepared in M4, will be updated in M18 and M36. The DMP provides a dynamic working framework for ongoing improvements.

The PROTEIN4IMPACT consortium is committed to transparency and open science principles while balancing ethical, legal, and commercial considerations.

2   Protein4Impact Project Summary

PROTEIN4IMPACT project is dedicated to assessing the nutritional, health, safety, and quality characteristics of novel protein foods (NFs) derived from unconventional sources, while also evaluating their environmental and socio-economic implications. To ensure a holistic approach and maximize the project's contribution to the transition toward healthier protein-based foods, PROTEIN4IMPACT focuses on utilizing selected agri-food, fisheries, and aquaculture by-products as primary raw materials. Additionally, natural protein producers—including fungi, bacteria, insects, and both micro- and macroalgae—play a key role in this alternative protein production process. Secondary by-products generated along the value chain will be repurposed for reuse within the food production cycle, materials or as a source of process energy.

The extracted proteins will undergo modification and functional enhancement, followed by sensory evaluation and testing in human food trials, as well as in aquaculture as feed components. To assess the feasibility of large-scale NF production, the project will simulate industrial-scale conditions using various methodologies and tools, including Life Cycle Assessment (LCA), Life Cycle Costing (LCC), Risk Assessment (RA), Techno-Economic Analysis (TEA), and Acceptability Optimization (AO), along with Digital Twins, real-time AI-driven monitoring, and Decision Support Systems (DSS).

At the final stage, key parameters influencing protein NF production will be optimized to enhance protein content, while impact assessments will be carried out at each phase—extraction, characterization, modification/functionalization of protein-rich fractions, and NF production—as well as across the entire integrated process. The project will also evaluate consumer acceptance and the societal relevance of NFs through sensory testing in various markets, allowing for a comparative analysis between Europe and other global regions.

By adopting the comprehensive, interdisciplinary approach proposed in PROTEIN4IMPACT, the project aims to establish a realistic and sustainable framework for the effective use of alternative proteins in the food industry.

3   Protein4Impact Data Summary

3.1   Principles of the Data Management

PROTEIN4IMPACT is a Research and Innovation Action initiative funded under the Horizon Europe programme. Over its 36-month duration, the project will generate a wide range of digital datasets from various consortium partners. Establishing a robust Data Management Plan (DMP) is a crucial initial step to ensure effective data storage, accessibility, availability, and reuse, all in alignment with the FAIR principles.

The EU Open Science policy, as outlined in Annex V of the EU Grants Annotated Grant Agreement, alongside the indicative DMP template, has provided a structured foundation for designing a comprehensive strategy for data management within PROTEIN4IMPACT.

The Data Management Plan of PROTEIN4IMPACT is designed to:

i.   Develop and maintain an up-to-date DMP, with scheduled revisions at project milestones M18 and M36.
ii.   Handle all data in accordance with FAIR principles, ensuring it is easily findable, accessible, interoperable, and reusable.
iii.   Store research data in a trusted repository (Zenodo), making it publicly available as soon as possible following scientific publication.
iv.   Guarantee open access to research data, applying the latest versions of the Creative Commons Attribution International Public License (CC BY) or the Creative Commons Public Domain Dedication (CC0). Exceptions will be granted where openness could compromise the legitimate interests of beneficiaries or conflict with obligations under the Grant Agreement (refer to Article 17) or EU competitive considerations.
v.   Provide repository-based metadata and supporting information on research outputs, including any tools or methodologies needed for data reuse and validation.
vi.   Ensure open metadata (under CC0 licensing), capturing essential details such as dataset origin, publication details (authors, title, publication date, and venue), Horizon Europe grant identifiers (project name, acronym, and number), licensing terms, and persistent identifiers for publications, authors, organizations, and grants. Where relevant, metadata will also include persistent identifiers for related research outputs or tools necessary for validating the publication's conclusions.

This structured approach of the DMP will facilitate responsible data management, ensuring transparency, accessibility, and compliance with Open Science principles throughout the PROTEIN4IMPACT project.

3.2   Objectives of the Data Management Plan

The present deliverable, D1.1, constitutes the initial version of the PROTEIN4IMPACT Data Management Plan (DMP). This plan is designed to support the project team in effectively implementing research activities while ensuring compliance with the Grant Agreement and Open Science policy. Furthermore, it serves as a key tool for facilitating communication, disseminating project results, and fulfilling reporting obligations. This PROTEIN4IMPACT DMP has been developed as part of Work Package 1 (WP1), Task 1.1, with the following core objectives:

Identify and classify key research datasets generated throughout the project, linking each dataset to its corresponding research objective.

⇒   Ensure adherence to the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) and encourage Open Access and Open Data sharing, while considering necessary restrictions.

⇒ Define essential dataset details, including data formats, storage locations, access limitations for sensitive datasets (e.g., restricted access rights), and the selection of appropriate repositories for long-term storage.

⇒ Establish a flexible and evolving framework that allows continuous updates and improvements in data management practices over the course of the project.

⇒ Ensure compliance with ethical, legal, and security requirements, particularly regarding data protection, confidentiality, and intellectual property rights.

⇒ Implement clear guidelines for data documentation and metadata standards, ensuring datasets are well-described and easily interpretable for reuse and validation by external stakeholders.

⇒ Promote interoperability and standardization, ensuring that datasets are stored in widely accepted formats and are compatible with existing tools and platforms.

⇒ Facilitate data preservation and long-term accessibility, ensuring that project outputs remain available beyond the project's completion.

The PROTEIN4IMPACT DMP covers both digital research data and other research outputs, particularly public deliverables. It has been developed based on dataset information collected through an internal survey of PROTEIN4IMPACT consortium members, conducted between M3 and M4 of the project. This document will be revised and expanded at M18 and M36, ensuring that data management strategies evolve in response to project developments and emerging requirements.

3.3 Data Management in the project

Data management is a key task within WP1 (Project Management and Coordination), which is led by CS, the project coordinator. Within WP1, Task 1.1, focused on data management, is led by GG. The DMP form was prepared in coordination with the project coordinator and in close collaboration with all project team members, following a DMP template provided by EU authorities (EU Funding & Tender Portal - Reference Documents). This effort aimed to establish a solid foundation for the project's DMP, including its principles and objectives.

In addition, GG provides guidance to the entire project team throughout the data lifecycle and ensures they are informed about their relevant data management obligations. GG is also responsible for conducting periodic reviews of the DMP, working collaboratively with all project team members. All related questionnaires are securely stored internally on the project's share point.

3.4 Summary of PROTEIN4IMPACT data and other research outputs

The PROTEIN4IMPACT project is structured into seven work packages, five of them are dedicated to research and innovation. WP1 focuses on the dissemination, exploitation, and communication of project outcomes, as well as it is responsible for project management and coordination. WP7 deals with ethical aspects.

The PROTEIN4IMPACT project has identified broad spectrum of datasets that will be generated over its course. A brief description of these datasets, including their origin, format, and relation to the project's objectives, is provided in Table 1 *(See Attachments)*.

Additionally, the project will produce additional other research outputs, primarily public deliverables provided in Table 2 *(See Attachments)*.

3.4.1   Datasets

Table 1 presents a detailed summary of the datasets produced as part of the PROTEIN4IMPACT project, outlining their source, format, description, and significance in achieving the project's goals. These datasets primarily originate from laboratory experiments and encompass a broad spectrum of research activities carried out across multiple work packages and specific tasks within the project. They serve a crucial role in advancing the development of novel foods, evaluating their environmental performance, and constructing a comprehensive methodology database aligned with the project's core objectives.

The data collected span a diverse array of formats, tailored to different research needs. These include structured numerical data from analytical measurements, textual annotations providing experimental context, high-resolution images capturing microscopic and macroscopic observations. By integrating these varied data types, the project aims to create a robust scientific foundation for novel food technologies, fostering advancements in the novel food processing technologies.

3.4.2   Other research outputs

Table 2 provides a comprehensive overview of additional research outputs produced within the PROTEIN4IMPACT project, specifically focusing on non-public deliverables. These deliverables encompass a variety of formal project reports, which serve as key documentation of the research progress, findings, and methodologies applied throughout the project's duration.

Once submitted to the EC, their executive summaries will be made publicly available on the project's official website and, upon approval by the Project Officer, will also be archived within the PROTEIN4IMPACT community on Zenodo (https://zenodo.org/communities/PROTEIN4IMPACT), ensuring long-term accessibility and preservation.

Beyond formal reports, other research outputs covered in this table may include:

⇒ Scientific publications in peer-reviewed journals, detailing experimental methodologies, data analysis, and key discoveries.
⇒ Conference presentations and posters, showcasing project advancements at national and international scientific meetings.
⇒ Technical guidelines and best practice documents, providing recommendations for researchers and industry stakeholders on biosensor development and environmental assessments.
⇒ Training materials and educational resources, designed to support knowledge transfer, including workshops, webinars, and instructional videos.
⇒ Software tools and computational models, developed within the project for data analysis, biosensor performance prediction, and environmental impact assessment.
⇒ Policy briefs and stakeholder reports, summarizing key findings and their implications for industry, regulatory bodies, and decision-makers.

By consolidating these diverse research outputs, the PROTEIN4IMPACT project aims to maximize knowledge dissemination, promote collaboration, and support future advancements.

Overview in the Table 2 (See Attachments).

4 Protein4Impact FAIR data approach

4.1 <u>F</u>indable data including provisions for metadata

The PROTEIN4IMPACT project is built on a foundation of open collaboration and structured knowledge-sharing while strictly adhering to its data management policy. This policy prioritizes maximizing transparency and accessibility as early and extensively as possible while maintaining compliance with FAIR (Findable, Accessible, Interoperable, and Reusable) principles and the data management guidelines set by Horizon Europe (as specified in the Annotated Grant Agreement, particularly Article 17).

To enhance the discoverability and long-term accessibility of research data generated within the project, all datasets—along with other research outputs—will be assigned persistent identifiers (PIDs), preferably DOIs, through trusted repositories. The Zenodo platform has been designated as the primary archival repository, where the project has already established a dedicated PROTEIN4IMPACT community to store and manage datasets and related research deliverables.

All relevant datasets, except those excluded to safeguard intellectual property rights or potential patent applications, will be accompanied by comprehensive metadata frameworks. These metadata records will be made openly accessible under the CC0 public domain dedication via trusted repositories, ensuring seamless data sharing. The metadata will include:

⇒ Dataset descriptions, submission dates, authorship, deposition venues, and embargo periods.
⇒ Horizon Europe funding details, including the project's name, acronym, and grant number.
⇒ Licensing terms defining data usage permissions.
⇒ Persistent identifiers for datasets, researchers, and affiliated institutions, where applicable.
⇒ Links to related publications and research outputs, enabling AI-powered data harvesting and integration into broader scientific ecosystems.

To further enhance searchability and interoperability, PROTEIN4IMPACT datasets will be systematically categorized using standardized keywords. A detailed breakdown of data findability strategies, including metadata requirements, can be found in Table 3 (*See Attachments*).

4.2 <u>A</u>ccessible data

As part of the **PROTEIN4IMPACT** project, an initial questionnaire was distributed among project partners to assess the accessibility of datasets generated throughout the research process. The feedback collected has helped shape a data management strategy that ensures **efficient storage, accessibility, and traceability of research outputs related to novel food innovations**.

All datasets produced within the project—whether **open or restricted**—will be deposited in **Zenodo**, a trusted open-access repository. This ensures that each dataset is assigned a **persistent identifier (DOI)**, making it easily citable and linkable to relevant scientific publications, supplementary datasets, or complementary research findings.

In addition to Zenodo, partners may opt out to use its **institutional repository**, as an alternative data storage platform. Each **dataset version** will be systematically tagged and accompanied by **comprehensive metadata**, with clearly defined **linkages between different dataset iterations**. To enhance dataset discoverability, the **project team will establish a standardized naming convention**, which includes:

- **A prefix** to specify whether the file is a dataset, metadata, or template.
- **A core identifier** composed of:
    - A concise, meaningful name describing the dataset or template.
    - The acronym or short name of the data-contributing partner(s).
- **A suffix** indicating the last upload date in **YYYYMMDD** format.

This structured approach ensures that datasets are **well-organized, easily searchable, and efficiently archived**. PROTEIN4IMPACT team members are strongly encouraged to utilize these repositories for data storage and archiving of project results.

Data Accessibility and Restrictions

Datasets supporting **scientific publications** will be **deposited at the time of publication**, ensuring immediate availability through **free and standardized access protocols**. While most datasets and research outputs will comply with **FAIR principles (Findable, Accessible, Interoperable, and Reusable)** and be **openly available**, some datasets will have **restricted access due to their potential for intellectual property protection and future patent applications**.

Long-Term Data Availability and Licensing

To **ensure longevity**, project datasets will remain **accessible for a minimum of five years** after the official completion of PROTEIN4IMPACT. **Publicly available datasets** will continue to be hosted on **Zenodo**, ensuring long-term preservation and accessibility.

- **Open-access datasets** will be licensed under **Creative Commons Attribution International (CC BY)**, allowing others to freely use and build upon the data while crediting the original authors.

- **Metadata** will be published under the **Creative Commons Public Domain Dedication (CC0)**, enabling unrestricted sharing and reuse.

Ethical, **legal, and intellectual property considerations** surrounding **data sharing and commercialization** will be carefully managed to **balance openness with innovation protection**, ensuring that the project's novel food research can be effectively leveraged by both academic and industry stakeholders.

4.3   Interoperable data

Interoperability is a fundamental principle within the PROTEIN4IMPACT project, ensuring that datasets related to novel food innovation can be seamlessly integrated, shared, and utilized across different systems, disciplines, and research initiatives. To maximize interoperability, the project adheres to structured data standardization practices, metadata frameworks, and persistent identifiers, enabling smooth data exchange among various research communities, stakeholders, and digital platforms.

All datasets produced within the project—whether openly accessible or restricted—will be deposited in Zenodo, a widely recognized open-access repository. Zenodo provides each dataset with a Digital Object Identifier (DOI), ensuring it remains permanently citable and traceable while being linked to relevant scientific publications, supplementary datasets, or complementary research findings. This approach facilitates cross-referencing and supports broader scientific engagement with the project's outputs.

Additionally, the partners may choose to complement Zenodo with their institutional repository, which allows for version control and enhanced dataset management. Within such institutional repository, each dataset version will be:

- Systematically tagged, ensuring clear differentiation between data iterations.

- Accompanied by comprehensive metadata, improving discoverability and contextual understanding.

- Linked to previous versions, enabling researchers to track dataset evolution over time.

- To further support data interoperability, the opting out team will establish a structured naming convention, ensuring consistency and clarity across different datasets. This convention consists of:

    o **A prefix** specifying whether the file represents a dataset, metadata, or a template.

    o **A core identifier**, composed of:

    o **A concise**, descriptive name of the dataset/template.

    o **The acronym** or short name of the data-contributing partner(s) (PFAStwin as the default for templates).

    o **A suffix** indicating the date of the last upload in YYYYMMDD format.

By implementing these interoperability measures, PROTEIN4IMPACT enhances the searchability, accessibility, and usability of its datasets, ensuring they can be easily integrated into broader scientific and industrial applications. Researchers, policymakers, and food innovators will be able to seamlessly access and analyze the data, fostering cross-sector collaboration in novel food development.

4.4 Data Re-use

One of the key objectives of the **PROTEIN4IMPACT** project is to ensure that research data generated within the framework of **novel food innovation** is not only accessible but also effectively **reusable** by diverse stakeholders, including **scientists, industry partners, regulatory bodies, and policymakers**. Reusability is crucial for **accelerating innovation, reducing redundant research efforts, and enhancing scientific transparency**.

To facilitate reusability, datasets supporting **scientific publications** will be **deposited at the time of publication**, ensuring immediate availability through **free and standardized access protocols**. The majority of research outputs align with **FAIR principles (Findable, Accessible, Interoperable, and Reusable)** and are openly available. However, certain datasets with potential for **patent applications and commercialization** will have restricted access.

Additionally, the project promotes **responsible data-sharing practices** by employing:

- **Creative Commons Attribution International (CC BY) licenses** for open-access datasets, allowing others to **reuse, modify, and redistribute** the data, provided that proper credit is given.

- **Creative Commons Public Domain Dedication (CC0) for metadata**, ensuring unrestricted availability and enabling **metadata harvesting by AI-driven tools** for enhanced searchability and integration.

Ethical, **legal, and commercial considerations** will be carefully managed to **balance open science principles with the need for intellectual property protection**. By enabling **broad data reusability**, the PROTEIN4IMPACT project supports **scientific progress, regulatory advancements, and industrial applications**, contributing to the development of **next-generation sustainable food solutions.**

5 Conclusion

The Data Management Plan (DMP) for the PROTEIN4IMPACT project has been designed to ensure that all research data generated throughout the project's lifecycle is handled in a responsible, transparent, and sustainable manner. By adhering to the FAIR principles (Findable, Accessible, Interoperable, and Reusable) and complying with Horizon Europe requirements, the project guarantees that its datasets contribute meaningfully to scientific progress, regulatory frameworks, and industrial innovation in novel food development.

Through the Zenodo repository, supplemented eventually by institutional archives, PROTEIN4IMPACT ensures that all datasets are securely stored, well-documented, and easily retrievable. The use of persistent identifiers (DOIs), standardized metadata frameworks, and structured naming conventions enhances interoperability and data discoverability, facilitating seamless integration across academic, industrial, and policy-making sectors.

The project actively promotes open data practices, making the majority of its datasets publicly accessible while responsibly managing intellectual property and ethical considerations for restricted data. Long-term preservation strategies guarantee that research outputs remain available for at least five years after project completion, ensuring ongoing impact and reusability in future food innovation initiatives.

By implementing this comprehensive Data Management Plan, PROTEIN4IMPACT strengthens collaborative research efforts, accelerates scientific discoveries, and supports sustainable food solutions. The project's commitment to data openness, integrity, and usability will contribute to a more transparent and efficient research ecosystem, ultimately benefiting academia, industry, and society at large.

5. Attachments

Table 1: PROTEIN4IMPACT project datasets – description

Table 2: Other research outputs – project deliverables

Table 3: Findability of data including provisions for metadata (tba in M18 update)

Table 4: Interoperability of data (tba in M18 update)

Table 5: Reusability of data (tba in M18 update)

Table 1: PROTEIN4IMPACT project datasets – description

Dataset No	Name of the dataset	WP/Task	Leading partner	Data description
ER 1.3	Project Handbook	WP1	CS	Gathers all information in one document. Provides detailed guidance, management processes, and procedures for effective project execution and collaboration.
ER 1.4	List of past and ongoing EU projects	WP1	CS	Made to contact each coordinator for encouraging the establishment of a collaborative link
ER 1.5	Mapping and database of the relevant stakeholders	WP1	CS	Guide engagement strategies, ensure relevant involvement, and maximize project impact.
ER 1.6	Platform and dietary data hub sharing	WP1	IDENER	Creation of a centralized digital repository that not only stores all official consortium documents in one place but also enables seamless online collaboration. This system ensures structured, easy, and immediate access to essential documents for all partners. Additionally, it fosters real-time teamwork, streamlining workflows and improving efficiency in document management and decision-making.
ER 2.2	Bacterial fermentation process for protein from waste	WP2	DTU	Membrane aerated bioreactor for safe and resource efficient aerobic fermentation of hydrogenotrophic bacteria for microbial protein
ER 2.3	Elaboration of experimental diets for rearing larvae	WP2	ENEA	Rearing larvae using agricultural by-products represent a way to reduce waste and provide valuable source of protein
ER 2.4	Collection fish by-products samples and defatting process	WP2	CNR	The use of fish by-products as valuable source of protein is a way to reduce waste production
ER 2.5	Solid fermentation bioprocess method	WP2	ENEA	Solid fermentation bioprocess to obtain food-grade mycoprotein from food industry by-products
ER 2.6	Protein concentrates obtained by extrusion	WP2	CS	Using extrusion to modify biomass in order to recuperate the protein fraction
ER 2.7	Protein derived from macro- and microalgae	WP2	ENEA	Macro- and microalgae protein improvement
ER 3.2	alternative proteins (soy, algae) with improved	WP3	UOH	Document describing fermentation of soy proteins and algae with basidiomycetes

	aroma and technofuctionality			for improved sensorial attributes and technofunctionality
ER 3.5	Human acceptability of alternative proteins	WP3	UOH	Information on sensory profiles and human acceptability of several alternative proteins from WP2 and 3
ER 4.1	Online service for New Foods protein production simulations using AI Digital Twin of realistic industrial infrastructure of typical food processors	WP4	GOLEM	Innovation online digital service for simulation of different protein production and sales scenarios and in-depth holistic assessment of industrial results and impacts by stakeholders using automated dashboards, interactive analytics reports.
ER 4.2	Open cyber-physical knowledgebase model of realistic food processor enterprise	WP4	GOLEM	The open model will allow easy customisation to optional protein production scenarios, usage of various KPIs and sustainability conditions to provide detailed insights into different approaches and conditions, including financial viability, quality management and environmental impacts
ER 5.2	Behavioral intervention methodology	WP5	UNIWA	A systematic review will be conducted providing an up-to-date in-depth overview of drivers of acceptance of a wide range of alternative proteins.
ER 6.1	Environmental impact assessment report	WP6	AgriClima	Environmental impact assessments complete and disseminated within stakeholder group, with final LCA and LCC.
ER 6.2	Risk assessment report	WP6	IDENER	Planning, monitoring, and execution of the environmental, health, and safety Risk Assessment (RA) and TEA.
ER 6.3	Recommendations for executing an SSbD framework	WP6	In.BIO	Project overall impact optimized, results shared broadly, exploitation plans in place, and recommendations for executing an SSbD framework disseminated

Table 2: Other research outputs – project deliverables

Dataset No	Name of the dataset	WP/Task	Leading partner	Data description
ER 1.1	DMP Plan (D1.1)	WP1	GG	Strategic plan for the data management, with the aim to fulfill the KPI's
ER 1.2	PEDR Plan (D1.2)	WP1	GG	Strategic plan for the project DEC activities, with the aim to fulfill the KPI's
ER 1.7	Practice abstract batch (D1.3)	WP1	CS	Highlights research outcomes' practical applications, exploitation potential, and sectorial impact for stakeholders.
ER 1.8	Policy brief (D1.8)	WP1	CS	Document outlining key recommendations, findings, and implications related to the production, extraction, transformation and commercialization of alternative proteins and products containing alternative proteins
ER 2.1	Report on IMPROVE alternative protein production (D2.1)	WP2	ENEA	This dataset constitutes a mandatory project deliverable
ER 3.1	Modified, funcionalized and characterized protein for their conversion into food and feed (D3.1)	WP3	UOH	Document describing all enzymatic treatments performed on the alternative proteins as well as the characterization of the enzymatic modified protein
ER 3.3	Report on pilot scale production process of selected compounds. Feasibility study (D3.3)	WP3	VRLS	This dataset constitutes a mandatory project deliverable
ER 3.4	Quality assessment of novel food produced with alternative proteins (D3.4)	WP3	NTUA	This dataset constitutes a mandatory project deliverable
ER 3.6	Report on testing alternative proteins as fish feeds (D3.6)	WP3	AQB	This dataset constitutes a mandatory project deliverable
ER 4.3	Technology platform to model protein production plants (D4.1)	WP4	GOLEM	This research output is a mandatory project deliverable
ER 4.4	Digital Twin simulator (D4.3)	WP4	GOLEM	This research output is a mandatory project deliverable

ER 5.1	Report on citizen satisfaction survey (D5.1)	WP5	UNIWA	It would be the first time that these novel foods (NFs) will be consumed by customers thus their feedback and satisfaction will be indicative for success of the project
ER 5.3	Report on data collection related to the drivers of acceptance of unconventional proteins (D5.2)	WP5	NDF	This dataset constitutes a mandatory project deliverable
ER 5.4	Report on agrifood stakeholders training activities (D5.3)	WP5	In.Bio	This dataset constitutes a mandatory project deliverable
ER 6.4	Report on energy recovery from production processes (D6.1)	WP6	In.BIO	This dataset constitutes a mandatory project deliverable
ER 6.5	Financial and environmental sustainability (LCC/LCA) and DSS (D6.2)	WP6	AgC	An innovative ex-post assessment will be used to evaluate the financial and environmental sustainability of the proposed solutions over the short and long term, using modern LCA and costing (LCA/C) theories.
ER 6.6	EOSC, dietary guidelines and regulatory framework (D6.3)	WP6	IDENER	IMPROVE data integrated into EOSC and a comparative analysis of current European dietary guidelines and consumption habits performed.
ER 7.1	OEI - Requirement No. 1 (D7.1) - MoU, Conflict of interest, CV of the Ethics Advisor	WP7	CS	Document detailing the mechanisms that prevent conflicts of interest, maintaining scientific integrity, and ensuring the credibility of research findings; document detailing the role of the Ethics Advisor; Curriculum vitae of the Ethics Advisor
ER 7.2	OEI - Requirement No. 2 (7.2) - Report 1	WP7	CS	Analysis of the activities conducted for the project from an ethics point of view, mainly to ensure compliance with EU ethical standards.
ER 7.3	OEI - Requirement No. 3 (D7.3) - Report 2	WP7	CS	Analysis of the activities conducted for the project from an ethics point of view, mainly to ensure compliance with EU ethical standards.

Table 3: Findability of data including provisions for metadata (tba in M18 screening)

Dataset No.	Will the dataset be identified by a persistent identifier?	What metadata will be created?	Metadata in a trusted repository under OA(CC0)	In which trusted repository do you plan to publish the metadata on your dataset?

Table 4: Interoperability of data (tba in M18 screening)

Dataset No.	Standards, formats or methodologies	Ontologies or vocabularies	Data format(s)	Software tools needed to access the data	Reference to other data

Table 4: Interoperability of data (tba in M18 screening)

Table 5: Reusability of Data (tba in M18 screening)

Dataset No.	Validation of data analysis and facilitation of data re-use	Data availability in the public domain to permit the widest re-use possible?	Relevant data quality assurance processes

6. List of Figures

n/a

7. List of Tables

n/a

8. List of Abbreviations

Abbreviation	Abbreviation for
CC0	Creative Commons Public Domain Dedication
CC BY	Attribution International Public License
DMP	Data Management Plan
OA	Open Access

9. Project Consortium

https://www.centralesupelec.fr

https://www.dtu.dk

https://katanalotis.org.cy

https://www.uni-hohenheim.de

http://www.enea.it

https://nordicdf.eu

https://www.venusroses-labsolutions.eu

https://www.cartif.es

http://www.grant-garant.cz

https://www.uniwa.gr

https://golem.at

https://www.cnr.it

https://www.ntua.gr

https://www.aquabt.com

http://www.agriclima.eu/

https://www.uminho.pt

http://www.idener.es

http://www.inbio.it