



Funded by
the European Union

Data Management Plan and IPR Strategy

UNIVERSITY OF VIENNA
30. March 2023

D8.1

Data Management Plan and IPR Strategy

Document ID D8.1

Due date 31/03/23

Submission 30/03/2023

Dissemination Public

Work Package 8

Author(s) Tima Anwana, Nikolaus Forgó,
Katja Hartl, Marie-Catherine Wagner

**Document
Version** 0.2

Grant Agreement 101059813

Duration 48 Months

Start Date 1 October 2022

End Date 30 September 2026



Funded by
the European Union

Contributors

Name	Organisation
Tima Anwana Nikolaus Forgó Katja Hartl Marie-Catherine Wagner	UNIVIE
Representatives of all Partners	All Partners

Revision history

Version	Date	Reviewer	Modifications
V0.1	22/03/2023	WR data manager	Draft reviewed and suggestions made
V0.2	30/03/2023	Ine van der Fels-Klerx	Final draft checked, and edits made

Disclaimer: The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained herein.

Index of Contents

1 Executive Summary.....	7
2 Introduction.....	9
2.1 The HOLiFOOD Project.....	9
2.2 Objectives of this Data Management Plan.....	9
2.3 Intended Audience	10
2.4 Data Collection.....	10
2.5 Updates	11
3 HOLiFOOD Data Lifecycle	12
3.1 Collecting Data	13
3.2 Processing and Analysing Data.....	14
3.3 Publishing and Sharing Data	14
3.4 Preserving Data	15
3.5 Re-Using Data	16
4 FAIR HOLiFOOD Data	18
4.1 Findable.....	18
4.2 Accessible	19
4.3 Interoperability of Data	20
4.4 Resuse of Data	21
5 FAIR HOLiFOOD Data Sets	23
5.1 Administrative Data	23
5.2 Stakeholder Data.....	25
5.3 Data to be integrated in the HOLiFOOD Platform.....	27
5.4 Software Data for the Platform.....	33
6 Data Protection and Ethical Aspects.....	37
6.1 Personal Data Collection.....	37
6.2 Informed Consent for Involving Stakeholders.....	37
6.3 The HOLiFOOD Website Privacy Policy	38
6.4 Transfer of Data in a Thrid Country	39
6.5 Ethical Considerations	39

7 IPR Strategy.....	42
7.1 Ownership of Project Results	42
7.2 Background Knowledge and Innovations.....	43
7.3 Rights of Access.....	44
7.4 Transfer of Rights	48
7.5 Steps to Ensure Compliance with the IPR Strategy	49
8 Conclusion	50

Index of Tables

<i>Table 1: FAIR Management Data</i>
<i>Table 2: FAIR Stakeholder data</i>
<i>Table 3: FAIR Platform Data</i>
<i>Table 4: FAIR Software Data</i>
<i>Table 5: Description of Background</i>
<i>Table 6: Access Rights for Implementation Purposes</i>
<i>Table 7: Access Rights for the Exploitation of Results</i>

Index of Figures

<i>Figure 1: Lifecycle of data during a research process</i>
<i>Figure 2: FAIR Principles</i>
<i>Figure 3: Data flow of the AGINFRA+ data linking workflow</i>
<i>Figure 4: HOLiFOOD Platform Infrastructure</i>

Acronyms List

CA	Consortium Agreement
DMP	Data Management Plan
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable and Re-usable
GA	Grant Agreement
GDPR	General Data Protection Regulation
HOLiFOOD	Holistic approach for tackling food systems risks in a changing global environment
IPR	Intellectual Property Rights
MS	Microsoft
UN	United Nations
WP	Work Package

1. Executive Summary

The EU funded project HOLiFOOD (Holistic approach for tackling food systems risks in a changing global environment /HORIZON-CL6-2021-FARM2FORK-01-15/) contributes to the UN's Sustainable Development Goals by improving the integrated food safety risk analysis (RA) framework in Europe. A major goal of this project is to promote the realisation of a truly secure and sustainable food production, whereby three supply chains are considered: cereals/maize, legumes/lentils and poultry/chicken. Applying a system approach, HOLiFOOD takes the whole environment into account in which food is being produced, and evaluates economic, environmental and social aspects. Artificial Intelligence (AI) and big data technologies are used to develop tools and methods for early warning systems and hazard detection. New holistic risk assessment methods and tools will be developed to support regulation in a changing global environment. This impact pathway is supported by an electronic data and knowledge sharing platform which aims at the full digitalisation of food (safety) systems and supporting transparency and impact for all stakeholders. In order to align with stakeholder priorities, preferences and user requirements, the HOLiFOOD innovations are designed and tested in a multi actor approach (i.e. Living Lab) involving all stakeholders (e.g., authorities, food producers and citizens).

In order to achieve these goals, various data types are produced and collected. Appropriate methods and tools for data collection and analysis have to be selected and a system for storing, organising, and protecting the data collected during the lifetime of the project must be developed. The HOLiFOOD project follows the Guidelines on FAIR Data Management in Horizon 2020¹. It aims to make data findable, accessible, interoperable and reuseable. At this stage of the HOLiFOOD project, the following datasets will be processed: administrative data, stakeholder data, software data and data to be processed on the HOLiFOOD platform. This DMP for HOLiFOOD data outlines the lifecycle of each identified dataset and describes how this data is made FAIR (findable, accessible, interoperable and fair). The DMP also aims to monitor privacy and confidentiality requirements regarding the generated data. Some key legal and ethical aspects, when it comes to

¹ H2020 Programme, 'Guidelines on FAIR Data Management in Horizon 2020'.
https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, July 2016.

processing of personal data, are presented in this document. However, a detailed analysis of the legal and ethical framework will be provided in D8.2 (M12). Furthermore, this DMP focuses on the management of project knowledge and ensures the fulfilment of the legal aspects of the partnership as outlined in the Consortium Agreement (CA) and Grant Agreement (GA).

2. Introduction

2.1 The HOLiFOOD Project

The overall objective of HOLiFOOD is to improve the integrated food safety risk analysis (RA) framework in Europe, in order to achieve the following objectives:

- i) Meeting future challenges arising from Green Deal policy driven transitions in particular in relation to climate driven changes;
- ii) Contribution to the UN's Sustainable Development Goals and
- iii) Support the realisation of a truly secure and sustainable food production.

HOLiFOOD will apply a system approach that takes into account the whole environment in which food is produced, including economic, environmental and social aspects. Three supply chains will be considered (i.e. cereals [maize], legumes [lentils] and poultry [chicken]). Artificial Intelligence (AI) and big data technologies will be used in the development of early warning and emerging risks prediction systems for known and unknown food safety hazards. In addition, tools and methods will be developed for hazard detection. Such methods will include both targeted and non-targeted approaches. Furthermore, new holistic risk assessment methods will be developed in which food safety risk will be embedded in a comprehensive cost-benefit analysis of the food system, including positive and negative health, environment and economic dimensions. An effective impact pathway will be developed and implemented by integrating the HOLiFOOD outputs into the legal framework, associated with the food risk analysis process. The impact pathway will be supported by an electronic data and knowledge sharing platform aiming at the full digitalisation of food (safety) systems and supporting transparency and impact for all stakeholders. In order to align with stakeholder priorities, preferences and user requirements, the HOLiFOOD innovations will be designed and tested in a multi actor approach (i.e. Living Lab) involving all stakeholders (e.g., authorities, food producers and citizens).

2.2 Objectives of this Data Management Plan

According to ISO 20294 data management is the “process of keeping track of all data and/or information related to the creation, production, distribution, storage, [...] use of e-media and

associated processes”². This Data Management Plan (DMP) outlines the main elements of HOLiFOOD data management policy. It is intended to cover the complete life cycle of the research data created and processed, and will show:

- i) The types of research data that will be generated or collected during the project;
- ii) How the research data will be processed and preserved;
- iii) Which parts of the datasets will be shared for verification or re-use;
- iv) The standards that will be implemented; and
- v) The handling of research data after the end of the project.

The DMP aims to monitor privacy and confidentiality requirements regarding the generated data. In addition, the DMP aims to ensure that the legal and ethical standards for data generation, use, storage, and share are applied throughout the project and in line with the overall-management of the project, as foreseen in the Grant Agreement and Consortium Agreement. Finally, the DMP seeks to ensure that appropriate technical standards are applied for data representation.

2.3 Intended Audience

This DMP is a public document. All persons interested in the topic will have access to this deliverable once it is published on the project website. However, it is assumed that this deliverable will be of primary interest to actors and stakeholders in the food industry and in the holistic risk management.

2.4 Data Collection

This iteration of the DMP is based on Information gathered by the HOLiFOOD Consortium through a questionnaire shared with all WP Leaders. The structure of the data management plan is based on the “Guidelines on FAIR Data Management in Horizon 2020”³. The DMP provides a current picture of the Project based on the answers from each WP.

² ISO 20294:2018, 3.5.4, data management, <https://www.iso.org>.

³ H2020 Programme, ‘Guidelines on FAIR Data Management in Horizon 2020’.

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, July 2016.

2.5 Updates

This document is the first iteration of the DMP (M6). The DMP will be updated and reviewed regularly in course of the project. Updated Versions will be submitted in M24 (D8.4) and M48 (D8.5).

3. HOLiFOOD Data Lifecycle

In the course of the HOLiFOOD project, a variety of data is produced, collected and processed. The planning stage of a research project is a crucial step for data management. Appropriate methods and tools for data collection and analysis have to be selected, and a system for storing, organising, and protecting the data collected during the lifetime of the project must be developed. This includes creating protocols for data collection, processing, and analysis. Another important aspect of planning data management of a research project is to plan consent for data sharing, which involves obtaining permission from study participants to share their data with other stakeholders. Finally, exploring existing data sources can help researchers to identify and reuse available data. By carefully planning each of these elements, researchers can ensure that their study is well-designed, ethical, and efficient.⁴

The proposal writing and the first months of the project are dedicated to planning the research and the data management process. The Grant Agreement outlines the research questions as well as the aims and goals of the project.⁵ Methods and tools for data processing and analysis were selected. This DMP gives a first overview of what kind of data are collected in the various WPs and how data are stored, organised and protected. Legal and ethical implications are identified and discussed with all partners (WP8). Appropriate steps are taken to ensure full legal and ethical compliance. UNIVIE as legal partner of the project provides guidance through the legal and ethical helpdesk.

An important aspect of data governance is a professional data life cycle management that monitors the project data from generation/collection onwards (Figure 1). The research data lifecycle describes the journey of data from its inception to its eventual disposal, and it helps

⁴ UK Data Service, 'Research Data Management' (UK Data Service) <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>.

⁵ Grant Agreement Project 101059813, Description of the action (DoA), Part A and B.

researchers understand the different stages of the research data process, including data collection, analysis, preservation, and sharing.

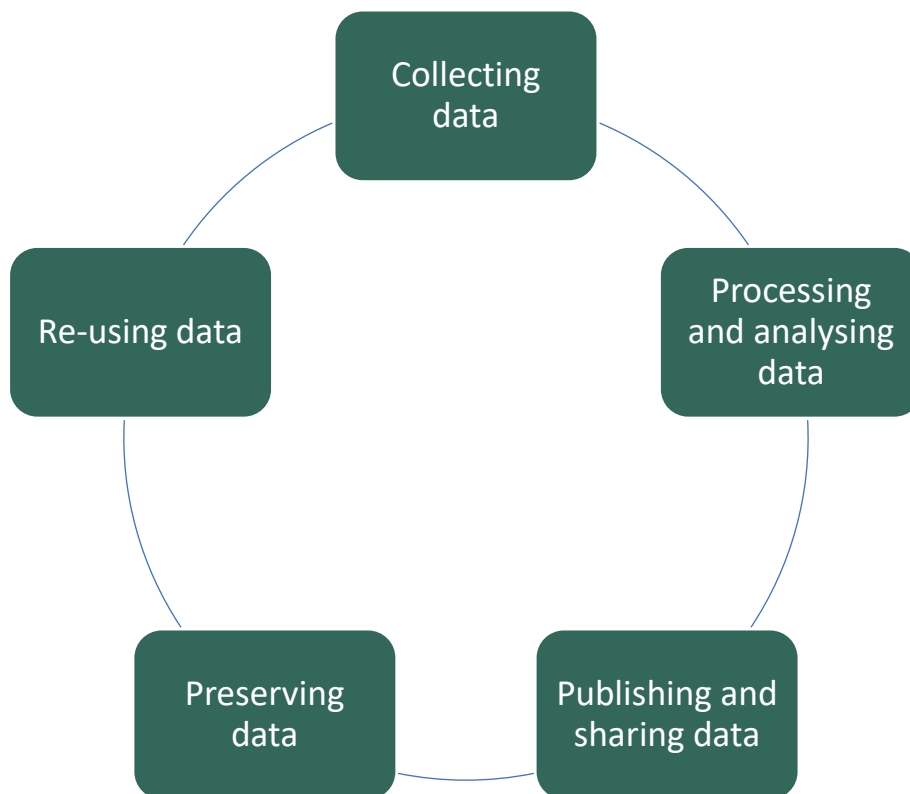


Figure 1: Lifecycle of data during a research process⁶

3.1 Collecting Data

Collecting data is the process of gathering information from various sources for the purpose of deriving and making informed decisions. The first step in data collection is to determine what data are needed and where they can be found.⁷

⁶ UK Data Service, 'Research Data Management' (UK Data Service) <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>.

⁷ UK Data Service, 'Research Data Management' (UK Data Service) <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>.

At this stage of the HOLiFOOD project, the following data categories are produced and collected:

- i) Administrative Data
- ii) Stakeholder Data (e.g. Living Labs)
- iii) Data to be processed on the HOLiFOOD Platform
- iv) Software data

3.2 Processing and Analysing Data

Data processing means transforming data into useful information. All data must be checked, cleaned and validated which means that data are correct and relevant. The quality of data is essential for a successful research project. According to ISO/IEC 25024 (2015), the quality of data is defined as the “degree to which the characteristics of data satisfy stated and implied needs when used under specified conditions”⁸. During the data analysis process, data analysis tools are applied in order to transform the data so that they can be easily interpreted, shared and published. The goal of data analysis is to discover useful information in order to draw informing conclusions, and support decision-making⁹.

This DMP outlines how each identified data set is being processed. Sections 4 and 5 of this DMP describe how data are being made findable, accessible, interoperable and reuseable (FAIR¹⁰) and how data are being protected.

3.3 Publishing and Sharing Data

Publishing and sharing data is an important step in making research more accessible and usable for a wider audience. Before publishing and sharing data, the following aspects should be considered:

- i) Establishment of an intellectual property rights (IPR) strategy.

⁸ <https://www.iso.org/obp/ui#search>

⁹ Cf. Transforming Unstructured Data into Useful Information", Big Data, Mining, and Analytics, Auerbach Publications, pp. 227–246, March 2014.

¹⁰ H2020 Programme, 'Guidelines on FAIR Data Management in Horizon 2020'.

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

- ii) Development of a documentation as guidance for the use of data and any associated tools.
- iii) Creation of discovery metadata to enable others to easily find and access the data. This can include information about the data, such as its format, scope, and data types.
- iv) Establishment of appropriate data access guidelines.¹¹

Once these steps have been taken, the data can be published or shared through various channels, such as online repositories or data sharing platforms. In the HOLiFOOD project, public deliverables will be openly accessible on the project website and are, therefore, open for re-use by all interested researchers and institutions. All partners in the consortium will engage in dissemination activities such as social media posts, webinars, papers and presentations to promote the project and its findings within the scientific community and the industry. The output will be made openly accessible whenever it is in alignment with the organising institution. Regarding scientific outcomes, the consortium is committed to the Horizon2020 Open Access¹² mandates and has embraced possible Open Access options known today. These include Gold Open Access, Green Open Access and self-archiving. The Consortium partners will therefore give preference to Open Access journals or non-Open Access journals that support Green and Gold roads.¹³

3.4 Preserving Data

Preservation of data is an essential component to ensure that valuable information can be accessed and used in the future, enabling continued research and innovation. To preserve data, it is important to maintain it to the best format, as technology changes over time and some file types may become obsolete. Data should also be stored and backed up in multiple locations to prevent loss. Preservation documentation should be created to provide guidance on how to preserve and maintain the data over time. Finally, data should be preserved and curated, with metadata added

¹¹ UK Data Service, 'Research Data Management' (UK Data Service) <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>

¹² H2020 Online Manual, 'Open Access & Data Management'. https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm.

¹³ H2020 Programme, 'Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data inn Horizon 2020'. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.

to provide context and ensure that the data is properly documented. Curating data involves making decisions about how data are stored, accessed, and shared over time, and can involve setting access policies and determining how the data are used.¹⁴

In the HOLiFOOD project, preserving data is of utmost importance. Wherever possible, common data formats are used to promote longterm interoperability. Wherever possible, specific standards such as DCAT metadata standards for publishing metadata of food records are used. Also standard classifications, such as FOODEX2 and AGROVOC, and standards ontologies for the country information are key to ensure interoperability. Data are stored locally on partner's infrastructure. Furthermore, HOLiFOOD uses Microsoft Teams¹⁵ provided by the Coordinator WR as a shared repository, where data are saved. Data are properly named and curated. During the collection and the pre-processing phase of each WP, the directory and file naming convention of the HOLiFOOD MS Teams repository is followed. Intuitive folder structure and a clear and unambiguous name is assigned at the folder and file name level. File naming contains the date, file type information and a version number. Metadata are also used. HOLiFOOD has also developed an IPR strategy. Section 7 of this deliverable outlines the Intellectual Property Rights (IPR) management strategy that all partners have assented to in the Grant Agreement and the Consortium Agreement.

3.5 Re-Using Data

Re-using data can offer a multitude of benefits in the field of research. Some of the most common ways to re-use data is by conducting secondary analysis and by undertaking follow-up research. Conducting research reviews can help evaluate the quality and validity of the existing research and provide a comprehensive understanding of the subject. The scrutinising and critical evaluation of findings can also help identify potential errors and inconsistencies that may require further analysis. Finally, re-using data for teaching and learning can provide a valuable resource for

¹⁴ UK Data Service, 'Research Data Management' (UK Data Service) <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>.

¹⁵ <https://www.microsoft.com/en-GB/servicesagreement/>

students and researchers for gaining a deeper understanding of research methodology and data analysis.¹⁶

The HOLiFOOD project builds upon existing data and also re-uses data. Data that are collected and produced are made available for re-use whenever it is possible. For this reason, a common Intellectual Property Rights (IPR) management strategy that all partners have assented to in the Grant Agreement and the Consortium Agreement was developed (see section 7 of this DMP). HOLiFOOD data may be useful to a range of actors including, researchers (e.g. researchers developing risk estimation and prediction models), stakeholders in the food industry, risk assessors and the general public.

¹⁶ UK Data Service, 'Research Data Management' (UK Data Service) <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>.

4. FAIR HOLiFOOD Data

According to the “Guidelines on FAIR Data Management in Horizon 2020”¹⁷, there are four principles (Figure 2) that should be applied in a Horizon 2020 project and in its Data Management Plan:

- i) Making data findable, including provisions of metadata,
- ii) Making data openly accessible,
- iii) Making data interoperable, and
- iv) Increasing data re-use.

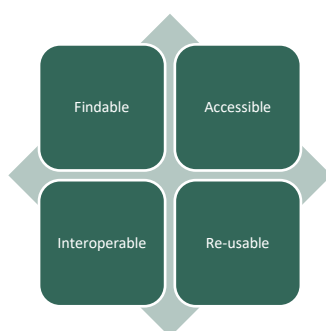


Figure 2: FAIR Principles

The following subsections outline what FAIR means and explain in brief how data in the HOLiFOOD project are made FAIR. Section 5 of this deliverable provides detailed information on how each identified dataset is made FAIR and being protected.

4.1 Findable

Data should be made **findable**. This requirement shall be fulfilled by outlining the discoverability of data, the identifiability of data and reference to standard identification mechanisms, the naming conventions used, the approach towards search keywords and clear versioning, as well as specification of standards for metadata creation.

¹⁷ H2020 Programme, ‘Guidelines on FAIR Data Management in Horizon 2020’.
https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, July 2016.

The HOLiFOOD project uses a MS Teams repository, which serves as a collective working space where the partners can exchange documents and collaborate in the development of deliverables. Intuitive folder structure and a clear and unambiguous name are assigned at the folder and file name level. Each set of data produced (dataset, deliverables, etc.) is named in a uniform way and includes a table with a version control. File names contain the date, file type information and a version number. For all deliverables, a standard template was shared among partners via the shared repository MS Teams. Deliverables follow the following naming convention: D[Deliverable Number] -[Deliverable Title]_ [Version of Deliverable].

In the context of platform data a “readme” file or other open format file is prepared per dataset as documentation, containing a list of the files in the dataset and information about these files; what they contain and how they relate. In some cases, the used program can integrate this information. Basic metadata on files generated are provided e.g., author, date created, date modified and file size. For sample data, where possible, information on the origin, date of harvest, etc. of the samples is also provided. Data will be available in a matrix form with the samples in rows and the analytical variables in columns. Data generated can be downloaded as excel-files. Metadata relevant to samples analysed for chemical contamination are provided according to the general EFSA Guidance on Standard Sample Description (SSD2) and Guidance on Data Exchange (GDE2). Wherever possible, specific standards are used, such as DCAT metadata standards for publishing metadata of food records, standard classifications for topics such as FOODEX2 and AGROVOC, standards ontologies for the country information and GS1 standards for publishing traceability information.

On the HOLiFOOD platform, all data entries inserted in the Platform will be assigned a unique identifier. Data are structured according to W3C open standards providing relevant identifiers and tags for querying. For software data, partners use the git repository naming convention: [component-name]-[subcomponent name]. Standard metadata provided by git repositories are used.

4.2 Accessible

The data should be made openly **accessible** whenever possible. Data accessibility is determined by specifying which data are made openly available, and if and why data is kept restricted. The

DMP elaborates on how the data are made available, what methods or software tools are needed to access the data, where the data and associated metadata, documentation and code are deposited as well as how access to data is provided in cases with restrictions.

All deliverables and reports will be published on the project website, as all deliverables in the HOLiFOOD project are marked as public in terms of the Grant Agreement. For all administrative data and data stored on the MS Teams HOLiFOOD repository, the project coordinator grants access to the repository to the partners actively involved in the project, in accordance with the Consortium Agreement and Grant Agreement. All personal data related to the stakeholders are restricted to consortium members. Only anonymised data will be made available to the public (in publications, events etc).

Parts of the data which will be integrated on the HOLiFOOD platform (developed in WP6) will be protected under access control. The correct execution of access will be ensured by the provided APIs and their documentation. This will be further assessed and developed at a later stage of the project. The results will be reported in the HOLiFOOD infrastructure & demonstrators deployment reports in WP6. Section 7 of this deliverable provides detailed information of the IPR strategy applied in the HOLiFOOD project. In accordance with the consortium agreement and the IPR strategy, some technical data are subject to access restrictions. However, technical data not subject to access restrictions will be made openly accessible, for example, open source software code will be accessible through the git repository and its URL.

4.3 Interoperability of Data

The **interoperability** of data should be ascertained by specifying what data and metadata vocabularies, standards or methodologies are used. In relation to this requirement, it should also be addressed whether standard vocabulary is used for all data types present in the datasets, as it may allow for inter-disciplinary interoperability.

HOLiFOOD aims to develop an open, distributed and interoperable platform. Partners will build upon existing technology services and platforms. HOLiFOOD will, for instance, build upon the work

of AGINFRA+¹⁸ (a former EU project). This project enables data exchange across different systems. AGINFRA+ has been a pioneer in data interoperability for food and agriculture and developed a core infrastructure service.

Whenever feasible, specific standards will be employed. For instance, DCAT metadata standards will be utilised to publish metadata of food records. Standard classifications such as FOODEX2 and AGROVOC will be used, and standards ontologies will be employed for country information. GS1 standards will be used to publish traceability information. Metadata will be recorded according to the general EFSA Guidance on Standard Sample Description (SSD2) and Guidance on Data Exchange (GDE2). The software code of the open-source software will be made available for reuse and review on GitHub, and open linked data will be published for open accessibility. To enable researchers to focus on specific areas of interest in their study, the open linked data is structured in accordance with open standards, and relevant identifiers and tags are used for querying.

4.4 Re-use of Data

The increase of data **re-use** should be documented by specifying how the data are licensed to permit the widest reuse possible. It will also be notified when and to whom the data will be made available for re-use, also after the end of the project, as well as how long the data will remain re-usable.

Data developed in the course of HOLiFOOD project may be useful to a range of actors including, researchers (e.g. researchers developing risk estimation and prediction models), stakeholders in the food industry, risk assessors and the public. Public deliverables will be openly accessible on the project website and as such are open for re-use by all interested researchers and institutions. Further, WP5 is dedicated to translate the outputs of WPs 1, 2, 3, 4 and 6 into societal use and policy practice. This is key to enhance the re-usability of the data produced in the project. WP5 ensures effective knowledge exchange with the public regarding emerging food risk identification.

¹⁸ Horizon 2020 Project 'Accelerating user-driven e-infrastructure innovation in Food & Agriculture'. <https://plus.aginfra.eu/>.

Administrative project data concerning management and finances is confidential within the consortium and not meant for re-use by third parties. All personal data collected in the course of the project will be kept confidential.

As stated above, the consortium partners will undertake various disseminating activities. WP7 is dedicated to project communication, dissemination and exploitation. HOLiFOOD aims to foster collaboration with other relevant projects and initiatives. Task 7.4 is dedicated to raise awareness of emerging food safety issues among consumers and citizens. Further, WP7 develops a strategy for a fast post-project exploitation. The consortium is dedicated to adhering to the Horizon 2020 Open Access mandates and has embraced various Open Access options, such as Gold Open Access, Green Open Access, and self-archiving. Consequently, the consortium partners will prioritise publishing in Open Access journals or non-Open Access journals that support the Green and Gold routes.

5. FAIR HOLiFOOD datasets

The tables below represent mappings of the lifecycle of each dataset produced and processed in the HOLiFOOD project following the FAIR principles.

5.1 Administrative Data

Administrative data (personal and non-personal data) are data generated or collected for organisational, managerial, transactional and record keeping purposes during the project. Some examples are:

- Partners' contacts
- Project planning data
- Templates of deliverables and reports
- E-mails and meeting minutes, documentation of communications among members of the projects
- Data from administration and financial management
- Data produced for dissemination purposes
- Deliverables and reports

Table 1: FAIR Management Data

DATA SUMMARY	
Data generated/collected	• Personal data: Partners' names, roles and contact details including email addresses. • Non-personal data: Project management data, communications between partners, meeting minutes, Deliverables and Reports.
Data Format	.csv; .xls; .xlsx, .doc, .docx, .pdf
Reproducibility	MS Teams and local repositories
Size of Data	Not defined at this stage.
Software tools for creating/processing/visualizing data	Microsoft Office, PDF Reader
Use of preexisting data – Data source	No
Storage and Backup Strategy	• MStTeams and local repositories

	Final and official submitted versions of deliverables are stored by the Project Coordinator on the European Commission Portal.
MAKING DATA FINDABLE	
Data Identifiers	Clear document naming such as deliverable number in the file name D[Deliverable Number] -[Deliverable Title]_ [Version of Deliverable].
Directory and file naming convention	- Each WP has an own folder on MS Teams - For all deliverables, standard templates are used. These templates were provided by partner EUFIC.
Search keywords for re-use	Deliverable number in the file name
Version numbering	Yes, such as v01, v02 or date
Metadata standards and automatic creation of metadata	no
MAKING DATA OPENLY ACCESSIBLE	
Open availability of produced data	- For the Consortium: MS Teams - For the Public: Public deliverables will be uploaded on the project website.
How will data be made accessible	- Via MStteams - Website
Methods or software tools and software documentation needed to access the data	MS teams
Deposition of data and associated metadata, documentation and code	No
Restrictions on use and access	Only project members have access to all the documents via MS Teams
Information retained in order for the data to be read and interpreted in the project and in the future	No
Correct execution of the access process	Only people added to MS Teams can access the documents. The HOLiFOOD project coordinator (WR) is in charge of monitoring access to the repository. Access is limited to specific persons indicated by the project partners and those who are granted access by the project coordinator.
MAKING DATA INTEROPERABLE	
Interoperability of data produced	Microsoft Office

Data and metadata vocabularies, standards or methodologies to facilitate interoperability	Microsoft Office
MAKING DATA RE-USABLE	
Data sharing and re-use	- Only consortium partners use confidential project documents. Personal data and confidential project documents will not be re-used. - Public deliverables are published on the project website
Audience for reuse	In the case of confidential project documents, the audience for re-use is limited to Consortium Partners. In the case of publicly disseminated deliverables, the audience for re-use is the public, specifically actors in the food supply chain, policy makers and researcher.
Period of re-usability	Open, at least until the end of the project.
Restrictions on data re-use (licenses)	Only project members
DATA SECURITY, ARCHIVING AND PESERVATION	
Security measures	- The security of documents stored on MS Teams - Each partner has its own security measures in place.
Archiving of data and long-term preservation	All the administrative data will still be online and accessible in MS Teams. The exact storage period will be defined in accordance with the project partners.

5.2 Stakeholder Data

WP4 is dedicated to the involvement of stakeholders and codesign in living labs. These living labs (LL) aim to:

- Identify and monitor food safety risks;
- Risk assessment and acceptance of holistic risks;
- Co-design the HOLiFOOD platform.

For each LL, an LL manager and LL facilitator are appointed. The manager is responsible for the coordination of the LL activities. The facilitator facilitates discussions and knowledge exchange. The Living Labs Guidelines (D4.1) outline the methodological approach of the LL. For each LL round, a report is written (D4.2, D4.3, D4.4). Furthermore, WP2 and WP5 engage with end-users in focus group meetings and workshops.

Table 2: FAIR Stakeholder data

DATA SUMMARY	
Data generated/collected	• Name • Contact details • Organisation • Role and Job title • Geolocatable information • Demographic information (such as gender, age) • Inputs obtained by questionnaire and workshops
Data Format	.csv; .xls; .xlsx, .doc, .docx, .pdf
Reproducibility	• Partner's local repositories with restricted access • HOLiFOOD repository on Teams site with restricted access • Transparent and documented process for stakeholder elicitation
Size of Data	Not defined at this stage. However, the size of the documents will not be significant.
Software tools for creating/processing/visualizing data	Microsoft Office
Use of preexisting data – Data source	Yes: existing materials provided by the respective partner regarding the elicitation process
Storage and Backup Strategy	• Individual computer • Partner's SHAREPOINT with restricted access. For Stakeholder data no third party storage should be used. • HOLiFOOD repository on Teams site with restricted access
MAKING DATA FINDABLE	
Data Identifiers	In the HOLiFOOD MS Teams repository, each WP has its own folder.
Directory and file naming convention	In the HOLiFOOD MS Teams repository, each WP has its own folder.
Metadata standards and automatic creation of metadata	no
MAKING DATA OPENLY ACCESSIBLE	
Open availability of produced data	All personal data related to the Stakeholders are restricted to Consortium Members. Only anonymised data will be made available to the public (in publications, events etc).
How will data be made accessible	HOLiFOOD MS Team repository
Methods or software tools and software documentation needed to access the data	MS teams

Deposition of data and associated metadata, documentation and code	No
Restrictions on use and access	Only HOLiFOOD Consortium Members
Information retained in order for the data to be read and interpreted in the project and in the future	Only the outcome of the Stakeholder meeting will be retained. All personal data will be anonymised or deleted after 10 years after the end of the project in accordance with the Grant Agreement.
Correct execution of the access process	Yes, only project partners will have access. The project coordinator (WR) gives access to the HOLiFOOD repository upon request. Each partner is responsible to get the access to the HOLiFOOD repository and other shared spaces related to the HOLiFOOD project restricted, once people are no longer working for the project.
MAKING DATA INTEROPERABLE	
Interoperability of data produced	Microsoft Office
Data and metadata vocabularies, standards or methodologies to facilitate interoperability	Microsoft Office
MAKING DATA RE-USABLE	
Data sharing and re-use	HOLiFOOD consortium members
Audience for reuse	HOLiFOOD consortium members
Period of re-usability	5 Years after the end of the project
Restrictions on data re-use (licenses)	HOLiFOOD consortium members. The use of Stakeholder data is limited to purpose set in the Informed Consent provided by the respective Stakeholder.
DATA SECURITY, ARCHIVING AND PERSERVATION	
Security measures	- Each partner has its own security system to protect the data - MS Teams security system
Archiving of data and long-term preservation	- Each partner store the data on their local repositories - MS Teams storage - EC portal

5.3 Data to be integrated in the HOLiFOOD Platform

Platform data refers to a variety of data that are produced, collected and processed mainly in WPs 1, 2, and 3. These datasets (models and approaches) are integrated on the HOLiFOOD infrastructure, which is developed in WP6.

In WP1 drivers, indicators and user expectations are implemented in the AI based models and tools for ERI predication (T1.1) and ERI monitoring (T1.2). WP1 builds on existing EI systems in the food safety domain. Optimal performing algorithms will be integrated in the HOLiFOOD platform. Further, three new text mining filters - in relation too legumes, cereals and poultry - are developed (T1.3).

In WP2 methods are developed and validated for the identification and characterization of existing and emerging hazards, with the aim to anticipate and possibly mitigate their impacts. Task 2.8 is dedicated to integrate the tools and data developed in WP2 in the HOLiFOOD platform.

In WP3, methods and tools for holistic risk assessment are developed to support regulation in a changing food environment. Food safety risks will be embedded in a comprehensive cost-benefit analysis. The approach is validated against poultry, cereals and legumes supply-chains. The methods and tools are integrated in the HOLiFOOD platform.

Table 3: FAIR Platform Data

DATA SUMMARY	
Data generated/collected	The datasets will be produced in WP1-3 and will be integrated in the HOLiFOOD platform. The data types will be defined in accordance with the results from WP4, the living labs. They will build the basis for defining the user requirements. In order to develop and validate the models, algorithms and tools in WP1, 2 and 3 various data types such as • Food safety monitoring data • Analytical lab test results • Microbial culturing data • Sequence data • qPCR results, validation data • Mass spectrometric (MS) data • Mathematical / statistical models to analyse the data • Cost-benefit analysis data will be generated and collected.
Data Format	Fastq, csv/excel, txt, R, Raw, wiff, qsession, mpr, csv/excel, txt, word, ThermoScientific *.raw format, SPA, CSV, obj, pptx, xlsx, usp, jpeg, tif, png, FASTA The HOLiFOOD Platform will be able to support various file types, including: CSV files, Excel files, Word Docs, Google Spreadsheets, RDF triples, JSON. Other data types based on the requirements collected by the LLs will be considered.
Reproducibility	• Partner's local repositories with restricted access

	HOLiFOOD repository on Teams site with restricted access
Size of Data	The underlying HPC infrastructure and databases will be able to support up to 1TB of data.
Software tools for creating/processing/visualizing data	Microsoft Office, R, Different software packages, preferably opensource packages, Omnic, TQ Analyst, PowerPoint, Excel, SpectraGryph, Origin, SIMCA 17, Image Viewer, Strip test reader software, fluorimetric and photometric reader software – for data generation and visualization; , Excel, R, Python, Microsoft Azure ML studio
Use of preexisting data – Data source	HOLiFOOD builds on existing external open-source data and prediction models.
Storage and Backup Strategy	During the collection and pre-processing in each WP, will be stored on: - Individual computer - Partner's SHAREPOINT with restricted access - HOLiFOOD repository on Teams site with restricted access On the HOLiFOOD infrastructure the following strategy, which will be further developed in the course of the project, will be employed: - Identification of critical data and systems that need to be backed up. - Determination of the frequency at which backups should occur. This frequency will depend on the criticality of the data, the rate of data change and the maximum tolerable downtime. - Establishment of a backup schedule and rotating backup media. This can be done by using automated backup software and backup devices such as tape drive, hard disk or cloud storage. - Storage of backups off-site or in a remote location for disaster recovery. This can be done by storing backups in a fireproof safe or an off-site data center Regular testing and verification of the integrity of the backup data, by regularly restoring a small subset of data from the backup and compare it with the live data. - Implementation of a disaster recovery plan that outlines the steps to be taken in the event of a data loss. This plan should include procedures for restoring data and systems, as well as procedures for communicating with stakeholders.
MAKING DATA FINDABLE	
Data Identifiers	In the data collection and pre-processing phase of each WP, individual folders with recognizable titles, are used.

	On the HOLiFOOD Platform all data entries inserted in the Platform will be assigned a unique identifier. Open Linked Data is structured according to W3C open standards and equipped with relevant identifiers and tags for querying.
Directory and file naming convention	The directory and file naming convention of the HOLiFOOD MS Teams repository is followed in the data collection and pre-processing phase of each WP. Intuitive folder structure and a clear and unambiguous name is assigned to the folder and file name levels. The file names contain the date, file type information and a version number.
Metadata standards and automatic creation of metadata	A "readme" file or other open format file will be prepared per data set as documentation, which will contain a list of the files in the data set and information about these files; what they contain and how they relate. In some cases, the used program can integrate this information. Basic metadata on files will be provided e.g., author, date created, date modified and file size. Data will be available using standards/matrix with the samples in rows and the analytical variables in columns. Data generated can be downloaded as excel-files. Metadata relevant to samples analyzed for chemical contamination will be provided according to the general EFSA Guidance on Standard Sample Description (SSD2) and Guidance on Data Exchange (GDE2). Wherever possible, specific standards will be used, such as DCAT metadata standards for publishing metadata of food records, Standard classifications for topics such as FOODEX2 and AGROVOC, standards ontologies for the country information and GS1 standards for publishing traceability information.
MAKING DATA OPENLY ACCESSIBLE	
Open availability of produced data	Parts of the data can be protected under access control. This will be further assessed and developed at a later stage of the project. See section 7 on the IPR strategy.
How will data be made accessible	• Will be in gitlab • HOLiFOOD Platform will provide APIs to access the stored data assets. • Analytical results will also be made available to the scientific community and stakeholders via journal publications and national and international conferences (e.g., poster & oral presentations). The complete datasets will be provided in open access publications, with the following options: publishing ('gold' open access), self-archiving ('green' open access) or using the Open Research Europe (ORE) platform.

Methods or software tools and software documentation needed to access the data	On the HOLiFOOD Platform, the APIs will be accompanied by a swagger-UI documentation
Deposition of data and associated metadata, documentation and code	• Gitlab • The data will be stored in the HOLiFOOD Platform, which will include services for data & knowledge integration/publication, services for predictive models integration/publication, and a shared encrypted database with private/confidential data.
Restrictions on use and access	Parts of the data can be protected under access control and encryption mechanisms. This will be further assessed and developed at a later stage of the project. Moreover, tools and methods developed in HOLiFOOD may be protected under IPR. See section 7 of this deliverable for further details.
Information retained in order for the data to be read and interpreted in the project and in the future	Data is accompanied with documentation to support reuse
Correct execution of the access process	On the HOLiFOOD platform, the correct execution of access will be ensured by the provided APIs and their documentation
MAKING DATA INTEROPERABLE	
Interoperability of data produced	Using standard data formats. HOLiFOOD will build upon the work of AGINFRA+ (a former EU project). This project enables data exchange across different systems. AGINFRA+ has been a pioneer in data interoperability for food and agriculture and developed a core infrastructure service. See figure 3 below this table.
Data and metadata vocabularies, standards or methodologies to facilitate interoperability	• Use of FoodOn ontology and RASFF vocabularies • Standards such as DCAT metadata standards will be used for publishing metadata of food records • Standard classifications for topics such as FOODEX2 and AGROVOC • Standards ontologies for the country information • GS1 standards for publishing traceability information • Common European Research Information Format (CERIF) is used as standard for recording metadata. • Metadata are recorded according to the general EFSA Guidance on Standard Sample Description (SSD2) and Guidance on Data Exchange (GDE2)
MAKING DATA RE-USABLE	
Data sharing and re-use	• Data are machine readable • Data are accessible through APIs to make them re-usable in other settings

Audience for reuse	The target audiences of HOLiFOOD are mainly scientists, industry, food safety authorities, risk assessors, regulators, etc.
Period of re-usability	At least during the project and 5 years after the end of the project. However, a longer period is envisioned. This will be further defined in the course of the project.
Restrictions on data re-use (licenses)	See section 7 on the IPR strategy.
DATA SECURITY, ARCHIVING AND PERSERVATION	
Security measures	The Platform data, during the collection and pre-processing in each WP, are protected by • The partners; each of them has its own security system to protect the data • MS Teams security system The HOLiFOOD Platform, which will be developed in WP6, will provide the ability to encrypt particular data records and control their access. This will be further developed in the course of the project. In case of data breach: • <u>Step 1:</u> Isolate the data breach to prevent any further compromise of personal information. • <u>Step 2:</u> Assess the data breach by gathering the facts and evaluating the risks, including potential harm to affected individuals and, where possible, taking action to remediate any risk of harm. • <u>Step 3:</u> Notify individuals and the Commissioner if required. If the breach is an 'eligible data breach' under the NDB scheme, it may be mandatory for the entity to notify. • <u>Step 4:</u> Review the incident and consider what actions can be taken to prevent future breaches.
Archiving of data and long-term preservation	At least during the project and up to 10 years after the end of the project as defined in the Grant Agreement.

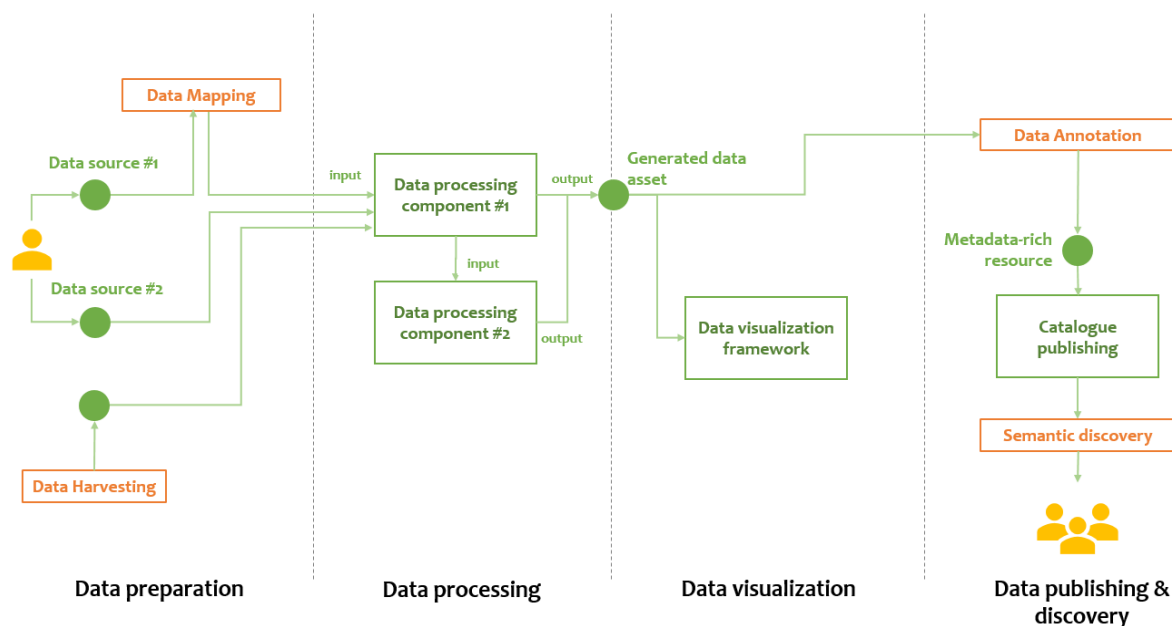


Figure 3: Data flow of the AGINFRA+¹⁹ data linking workflow

5.4 Software Data for the Platform

This category broadly refers to the main software development whose purpose is the design and implementation of an integrated European data and knowledge exchange platform. WP6 aims to design and develop the HOLiFOOD platform. Partners build on existing technology serviced and platforms. This allows to create an open, distributed and interoperable infrastructure (Figure 4). D6.1, D6.2, D6.3 and D6.4 are dedicated to describe the progress made in the development of the HOLiFOOD platform. The following partners are mainly involved: AGROKNOW, WR, CREME, and UVMB.

Table 4: FAIR Software Data

DATA SUMAMRY	
Data generated/collected	Software Code, Software Design, Platform Requirements

¹⁹ Horizon 2020 Project 'Accelerating user-driven e-infrastructure innovation in Food & Agriculture'. <https://plus.aginfra.eu/>.

Data Format	Java spring software code, Python code, SPARQL queries, RDF triples, JSON format, TTL format
Reproducibility	Software code will be stored and back-up strategies applied to ensure recovery, reproduction and reuse of technical data
Size of Data	The expected size will have to be estimated at a later point in the project and updated accordingly in subsequent iterations of this deliverable
Software tools for creating/processing/visualizing data	Decision support tools developed in T6.3, including appropriate end-user dashboards
Use of preexisting data – Data source	Pre-existing data are integrated from LOD (Linked Open data) Cloud, other EC funded research projects, partners systems and public/open data sources
Storage and Backup Strategy	The technical data (data sets, source code and docker containers) are stored and hosted into a private git server (Gitlab) and the HPC infrastructure of CREME.
MAKING DATA FINDABLE	
Data Identifiers	As provided by git repositories
Directory and file naming convention	The git repositories follow the naming convention: [component-name]-[subcomponent name]
Metadata standards and automatic creation of metadata	As provided by git repositories
MAKING DATA OPENLY ACCESSIBLE	
Open availability of produced data	Parts of the code will be made available open-source. In accordance with the IPR strategy outlined in section 7, this will further be defined at a later stage.
How will data be made accessible	The software code is accessible through the git repository and its URI.
Methods or software tools and software documentation needed to access the data	Access to software code is provided by standard git repository procedures and commands.
Deposition of data and associated metadata, documentation and code	Documentation is stored along the code in README.MD files using the standard structure of git repositories.
Restrictions on use and access	Parts of the code are confidential to ensure security and protect commercial exploitation and IPR.
Information retained in order for the data to be read and interpreted in the project and in the future	Software code is accessible through the git repositories, whose check-out URI will be made available.
Correct execution of the access process	Through standard git repository processes and commands.
MAKING DATA INTEROPERABLE	
Interoperability of data produced	Using standard data formats: Java spring software code, Python code, SPARQL queries, RDF triples, JSON format, TTL format

Data and metadata vocabularies, standards or methodologies to facilitate interoperability	Wherever possible, specific standards are used, such as DCAT metadata standards for publishing metadata of food records, Standard classifications for topics such as FOODEX2 and AGROVOC, Standards ontologies for the country information and GS1 standards for publishing traceability information.
MAKING DATA RE-USABLE	
Data sharing and re-use	The software code is structured into self-contained modules that can be re-used in other software projects
Audience for reuse	Software developers
Period of re-usability	Indefinitely, limited only by the potential obsolescence of the employed technologies
Restrictions on data re-use (licenses)	Parts of the code are protected by IPR and copyrighted by the contributing partners
DATA SECURITY, ARCHIVING AND PRESERVATION	
Security measures	The Platform will provide the ability to encrypt particular data records and control their access. In case of data breach: • <u>Step 1</u>: Isolate the data breach to prevent any further compromise of personal information. • <u>Step 2</u>: Assess the data breach by gathering the facts and evaluating the risks, including potential harm to affected individuals. Where possible, take action to remediate any risk of harm. • <u>Step 3</u>: Notify individuals and the Commissioner if required. If the breach is an 'eligible data breach' under the NDB scheme, it may be mandatory for the entity to notify. • <u>Step 4</u>: Review the incident and consider what actions can be taken to prevent future breaches.
Archiving of data and long-term preservation	The software code will be stored in the git repository indefinitely and at least for 5 years after the end of the project

WP6

WP1

Data

Models

WP4

Requirements

Feedback

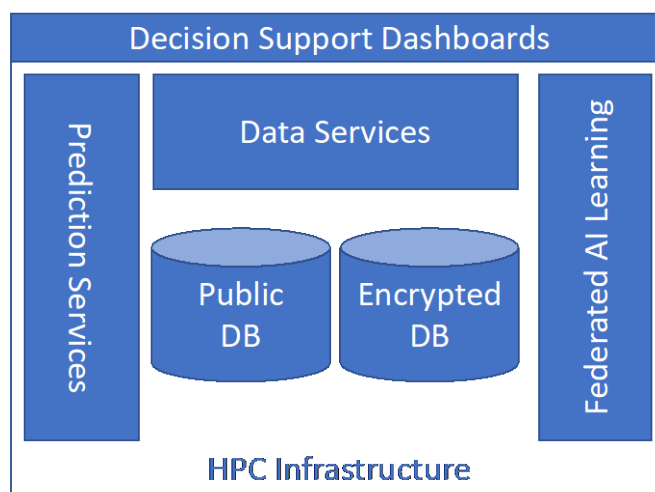


Figure 4: HOLiFOOD Platform Infrastructure

6. Data Protection and Ethical Aspects

This section outlines some key legal and ethical aspects, when it comes to processing of personal data. However, a detailed analysis of the legal and ethical framework will be provided in D8.2 (M12). As the project evolves, it will be updated in the next iteration of the framework report at the end of the project (D8.3). Furthermore, legal and compliance of the HOLiFOOD project is designed and ensured in WP8.

6.1 Personal Data Collection

Whenever personal data, information relating to an identified or identifiable natural person, are processed²⁰ special care has to be taken. Personal data are collected and processed in the following project related activities and tasks:

- i) Focus Group Meetings
- ii) Living Labs
- iii) Website

In the HOLiFOOD project, all data processing complies with EU law as well as national data laws and follows the guidelines on “Ethics and Data Protection”.²¹ Additional steps are to ensure a good balance between research objectives and the means used to achieve them.

6.2 Informed Consent for Involving Stakeholders

The involvement of external stakeholders is foreseen in various WPs. This participation requires legal and ethical considerations as well as appropriate legal agreements and documents. As part of its tasks under the legal and ethical helpdesk, UNIVIE has provided templates for legal documents that may be needed when involving Stakeholder data via the Legal and Ethical

²⁰ ‘Processing of personal data’ means any operation (or set of operations) performed on personal data, either manually or by automatic means such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction (Art 4(2) GDPR)

²¹ European Commission, Ethics and data protection, July 2021.

Helpdesk (email) in February 2023. These templates include information sheets and informed consent forms. This subsection outlines the main aspects of the informed consent used in the HOLiFOOD project. Deliverable 8.2 will provide a more detailed analysis and include the templates.

As the European Commission guidelines on ethics and data protection²² state, an informed consent is the “core of research ethics”.²³ Only after the appropriately prepared information has been conveyed to the research participants, and it has been fully understood, the body conducting research can look for and obtain the express permission to include the participants in the research. Using information sheets, HOLiFOOD partners, therefore, provide clear information about the project and what it expected from the participant. Consent will be sought after potential participants have had the possibility to ask questions.

Informed Consent is also required from a legal perspective and forms the legal basis for processing personal data of stakeholder. Article 4 (11) GDPR defines consent as a “freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her”. Each partner will ensure that individuals are fully informed about the ways in which their data will be collected, stored and used, and that they have given their consent to this process. In accordance with the GDPR, participants will be informed in clear language of the purpose and means of processing of their data, and of their rights, including their right to withdraw.

6.3 The HOLiFOOD Website Privacy Policy

Partner EUFIC is responsible for setting up the HOLiFOOD Website. In order to improve the user experience, HOLiFOOD will use Matomo Analytics Tool. Visitors of the HOLiFOOD website will be asked for their consent for the use of cookies. If the consent is not given, only technical cookies will be used. A detailed privacy policy drafted in collaboration with UNIVIE will be accessible on the

²² European Commission, *Ethics and data protection*, July 2021.

²³ European Commission, *Ethics and data protection*, July 2021.

website and regularly updated, this will be reported in the Legal and Ethical Framework Report (D8.2 – due in M12).

6.4 Transfer of Data in a Thrid Country

Two consortium partners, QUB and UNEW are located in countries, which are not part of the EU and the three additional EEA countries (Norway, Iceland, and Liechtenstein). In order to be able to share any data with these partners, all conditions set out in Chapter V of the GDPR (Article 44 to 50) must be fulfilled. Article 45 GDPR states that

“a transfer of personal data to a third country or an international organisation may take place where the Commission has decided that the third country [...] in question ensures an adequate level of protection. Such a transfer shall not require any specific authorization”.

This adequacy decision has the effect that personal data can be sent from an EU Member State or EEA state to a third country without any further safeguard being necessary. In the case of the UK, adequacy decision was adopted by the European Commission on 28 June 2021.²⁴ This adequacy decision is expected to last until 27 June 2025. In case personal data in the project may be sent from by the partners located in the EU member State or EEA state to the QUB and UNEW, no further safeguard is being necessary until this date. UNIVIE will further monitor the situation and keep partners updated if any changes occur.

6.5 Ethical Considerations

In order to achieve the best possible result for the project, it is important that some key ethical principles are taken into consideration. This first iteration of the DMP provides some main aspects, while the more detailed requirements for the following general rules will be described in the upcoming Deliverbale 8.2 (M12). Each partner contributing to the DMP is committed to promoting responsible research practices and to obtain the FAIR principle, which details can be found in section 4 in this document.

²⁴ European Commission, Commission Implementing Decision of 28.6.2021 pursuant to regulation (EU) 2016/679 to the European Parliament and of the Council on the adequate protection of personal data by the United Kingdom', July 2021.

HOLiFOOD follows EU rules and guidelines on FAIR Data Management and establishing trustworthy AI to achieve the aim of the project.²⁵ To ensure a high level of ethical standards, project partners are committed to follow the following measures regarding data management. Each WP Leader will monitor the activities of his/her WP. Further, the legal and ethical helpdesk will assist partners in ethical compliance issues.

- All partners check and verify the compliance and possibly required authorisation under national law for collecting and processing personal data. The general principle is that the consent of the persons concerned must be obtained.
- Participants have the opportunity to ask questions, receive training materials and engage in training sessions.
- Generally, for the ethical management of personal data, it is necessary to not only consider and mitigate risks for data subjects, but also to consider whether data processing can have a harmful impact (for example lead to discriminatory practices) towards groups of people who are not necessarily the data subjects.
- Non-personal data collected, processed and disseminated in the context of the project must be open to fair data exchange and equal access where appropriate.
- Infrastructures must be in place to enable secure storage and long-term archiving as well as research, access and re-use.
- Work Packages using pre-existing data must evaluate certain criteria such as methodology of data collection, accuracy, period of data collection, purpose and content of the data to guarantee high quality of the data.
- Personal as well as non-personal data such as business or software data must be protected with appropriate safeguards to protect against risks of loss, misuse or damage. This DMP provide information about how data are being handled and protected in the HOLiFOOD project.
- Data minimisation ensures that the collection of personal information is limited to what is directly necessary for the research purposes in the project.

²⁵ Grant Agreement Project 101059813, Description of the action (DoA), Part A and B.

- Results and finding of research will be presented honestly and openly. Clear and complete documentation of data editing, statistical procedures and assumptions will be provided.
- Open data can be freely used, modified, and shared by anyone for any purpose, while taking in consideration that not all publicly accessible data are open data. Partners and organisations will ensure access “to be as open as possible and closed as necessary”.²⁶ Where data (e.g. codes) will be protected by IPR, licences will be made available where necessary to ensure the best results possible.
- Creating and maintaining trust is one of the key challenges in the exchange of food safety data. The IPR strategy in HOLiFOOD (see section 7) aims to make the process as simple as possible, create trust among stakeholders, and prevent contractual imbalances between parties.
- Data can be biased if they are collected or analysed in a way that favours certain groups or individuals over others. This can lead to unfair treatment and discrimination. When personal data are gathered within HOLiFOOD, potential risks will be monitored within each Work Package, however the likelihood of unforeseen discrimination is considered to be low.
- WP8 developed a Legal and Ethical Helpdesk as a tool to offer guidance and support to the partners. A dedicated email address was established: holifood-helpdesk.id@univie.ac.at. Through the Helpdesk, partners are connected with legal researchers and experts at the University of Vienna (UNIVIE) to address any legal or ethical concerns. In addition to providing guidance, the Helpdesk also acts as a compliance monitoring tool,

²⁶ European Commission, Open Science, [Open Science \(europa.eu\)](https://open-science.europa.eu) (last accessed on 15.03.2023)

7. IPR Strategy

This section focuses on the management of project knowledge and ensures the fulfilment of the legal aspects of the partnership as outlined in the Consortium Agreement (CA) and Grant Agreement (GA). Furthermore, this section aims to monitor and ensure compliance with the CA and the Intellectual Property Rights (IPR) management strategy that all partners have assented to. In the context of the project, the CA and GA outline the IPR strategy, therefore, this section focuses on the governance of project knowledge, based on the provisions of the CA and GA.

7.1 Ownership of Project Results

The GA defines project results as “any tangible or intangible effect of the action, such as data, know-how or information, whatever its form or nature, whether or not it can be protected, as well as any rights attached to it, including intellectual property rights”.²⁷ Each project partner owns the results and innovations that they develop in the course of the project. Therefore, each partner retains the intellectual property rights in relation to their creations during the project.

In some instances, two or more beneficiaries may jointly own results or innovations. This happens in the following cases:

- i) When partners have jointly generated a result;
- ii) When it is not possible to establish the respective contribution of each beneficiary, or to separate them for the purpose of applying for, obtaining or maintaining their protection.

In the case of joint ownership, partners must enter into a Joint Ownership Agreement. This contract shall outline the terms of the Joint Ownership and ensure compliance with the obligations under the GA and the CA.

²⁷ Article 16.2, HOLIFOOD Grant Agreement.

In general, this IPR strategy does not cover all aspects pertaining to the collaboration between partners at the end of the project. The future sustainability of the project results will be determined by partners in the sustainability plan.

7.2 Background Knowledge and Innovations

All partners entered into the HOLiFOOD project with pre-existing knowledge and innovations that may be essential to the success of the project. This pre-existing knowledge is referred to in the GA and CA as “background”. Background refers to any knowledge that is held by a project partner prior to the commencement of the project and that is needed for the implementation of the project or the exploitation of the results.²⁸ Background may include data, expertise, information (tangible or intangible) and any rights such as intellectual property rights.²⁹ The partners retain the rights associated with all background that is used in the implementation of the project and the exploitation of project results. Table 5 outlines the existing background in the HOLiFOOD Project, in accordance with section 5 of the GA.

Table 5: Description of Background

Relevant Project Partners and Description of Background	
WFSR	• High resolution mass spectrometry libraries generated from standard compounds including mycotoxins, plant toxins and their modified forms • LC-HRMS analytical methods containing detailed inclusion and exclusion list or information dependent data acquisition • Metagenomics and generated sequence data
AGROKNOW	• Agroknow's Data Platform (ADP) • Food safety data from public authority sources, crawled and parsed by the ADP and made available through Riskadata API Service • Agroknow's Intelligence Stack (AIS) • FOODAKAI Software

²⁸ Article 16.1, HOLiFOOD Grant Agreement.

²⁹ Article 16.1, HOLiFOOD Grant Agreement.

BfR	• High resolution mass spectrometry libraries generated from standard compounds including mycotoxins, plant toxins and their modified forms • LC-HRMS analytical methods containing detailed inclusion and exclusion list or information dependent data acquisition • BfR NewsRadar
CREME	• DATA Science and Data Management Platforms: Data Foundry™ and Expert Models™ • Cloud computing infrastructure, access and version control, security, data models and structures, collaborative environment • Crème Global exposure and modelling software, including proprietary data and data engineering structures (Crème Food Safety, SAFE Project) • Machine learning, artificial intelligence and decision support systems (GLOPACK and others) • Know-how in the following domains: hazard, risk, food safety, migration modelling, exposure modelling, risk analysis, statistics, machine learning, data engineering, decision support systems, data analysis, data visualisation, scientific, reporting • Capabilities in UX design and software development
INRAE	• Food microbiology; performing challenge-tests on food matrices; culture of vegetative foodborne pathogens; gene expression quantification on foodborne pathogens (transcriptomics) • Predictive microbiology and statistics; QMRA, risk ranking; risk-benefit assessment; holistic assessment combining public health, environmental and economic impacts, multicriteria based on outranking methods. All above applied to food, food ingredients, food production, food supply-chain
WU	• Databases containing heat resistance and other characteristics of pathogenic and spoilage organisms (both used already in the scientific literature and not yet) for use in bench marking, model building, and quantitative assessments

7.3 Rights of Access

Partners retain IPR, including ownership rights of their background and the results they produce during the project. However, all partners are contractually required to grant access rights to other project partners. This section outlines the rules governing the provision and exercise of access rights in the project.

7.3.1 General Principles

The general rules regarding the granting of access rights are as follows:

- i) Access rights conferred by the CA and GA do not include rights to sublicense, unless this is expressly agreed upon in writing.
- ii) Access rights shall be granted free of any administrative fees.
- iii) Access rights are granted on a non-exclusive basis.
- iv) All requests for Access Rights shall be made in writing.
- v) The granting of Access Rights may be made conditional on the acceptance of specific conditions aimed at ensuring that these rights will be used only for the intended purpose and that appropriate confidentiality obligations are in place.
- vi) If a partner is no longer involved in the project, this does not affect its obligations to grant access.
- vii) If a partner defaults on its obligations, the other project partners may agree that that the defaulting partner no longer has access rights.

7.3.2 Access Rights for implementing the HOLiFOOD project

Article 9 of the CA and Article 16 of the GA outline the terms governing access rights to background knowledge. The respective partner who provides the background maintains all intellectual property rights associated. However, in terms of the GA all beneficiaries must be granted access to this background knowledge. Access must be granted on a “royalty-free basis” to background knowledge that is needed for beneficiaries to conduct their tasks in the project. Furthermore, these rights must be granted on fair and reasonable conditions. However, in certain instances, access to background knowledge may be restricted. These restrictions have been predetermined by partners and are highlighted in Table 6 below.

Table 6: Access Rights for Implementation Purposes

Relevant Partners	Background Knowledge	Access Rights for implementation purposes
WFSR	High resolution mass spectrometry libraries LC-HRMS analytical methods	Only granted to the extent that it is needed for implementation of the action

WFSR	Metagenomics and generated sequence data	Raw sequence data can be made available to project partners under embargo
AGROKNOW	Agroknow's Data Platform (ADP) Food safety data from public authority sources, crawled and parsed by the ADP and made available through Riskadata API Service	Copyrighted and protected by IPR Access rights granted project partners for the duration of the project and in the context of and for purposes of specific project tasks
AGROKNOW	Agroknow's Intelligence Stack (AIS) FOODAKAI Software	No access rights restrictions
BfR	High resolution mass spectrometry libraries LC-HRMS analytical methods	Only granted to the extent that it is needed for implementation of the action
BfR	BfR NewsRader	Accessible "as is". Access Right to is only granted to the extent that is needed for implementation of the action. Limited administrative support will be provided
CRÈME	See Table above	No access rights restrictions
INRAE	See Table above	No access rights restrictions
WU	Wageningen University databases	Access rights granted to project partners (given appropriate credits)

7.3.3 Access rights for the exploitation of results

Article 16 of the GA requires that partners must grant each other access to the “results needed for exploiting their results”. Therefore, partners must grant access to innovations that are developed in the project where this is necessary for the overall exploitation of project results. Requests for access may be made up to 12 months after the end of the project, unless partners agreed otherwise in writing.

Furthermore, partners must grant each other access to the background needed for exploiting their results. However, some background information is subject to certain access limitations for the purposes of exploiting the results of the project. These limitations are outlined in Table 7.

Table 7: Access Rights for the Exploitation of Results

Relevant Partners	Background Knowledge	Access Rights for exploiting the results
WFSR	High resolution mass spectrometry libraries LC-HRMS analytical methods	Prior consent of WFSR is required. Data generated outside the project by WFSR may only be used and published with permission of the WFSR
WFSR	Metagenomics and generated sequence data	Prior consent of WFSR is required.
AGROKNOW	Agroknow's Data Platform (ADP) Food safety data from public authority sources, crawled and parsed by the ADP and made available through Riskadata API Service.	Exploitation and access beyond the project duration is subject to prior negotiation between Agroknow IKE and the interested party. Written consent must be obtained from Agroknow IKE which must include specific terms for use and exploitation.
AGROKNOW	Agroknow's Intelligence Stack (AIS) FOODAKAI Software	No access rights restrictions specified
BfR	High resolution mass spectrometry libraries LC-HRMS analytical methods	Background may be used only with the prior consent of BfR. Data generated outside the project by BfR may only be used and published with permission of the BfR.
BfR	BfR NewsRader	prior consent of BfR is required
CRÈME	See Table 1 (above)	No access rights restrictions
INRAE	See Table 1 above	No access rights restrictions
WU	Wageningen University databases	Access rights granted to project partners (given appropriate credits)

7.3.4 Access Rights to Software

The general provisions outlined above, regarding access rights between project partners are applicable also to Software. However, according to the CA, access rights to Software does not include any right to receive source code or object code ported to a certain hardware platform. Furthermore, this right does not include any right to receive respective Software documentation in any particular form or detail, but only as available from the partner who grants the access right.

This section outlines the provisions governing rights of access between project partners during the development and deployment of the platform. It is important to note that external entities will be granted access to project results and use of the platform. The granting of rights to external entities is outlined in section 7.4 below and will be determined in the platform Terms of Use that will be reported in the legal and ethical framework deliverable (D8.2).

7.4 Transfer of Rights

7.4.1 Transfer of Ownership

In accordance with Article 8.3 of the CA, Each partner may transfer ownership, of its own results or innovations generated in the course of the project. However, the transfer of ownership to third parties and external entities shall not affect compliance of the GA or the CA. The beneficiaries must ensure that their obligations under the Agreements regarding their results are passed on to the new owner and passed in any subsequent transfer.

Article 8.3.2 of the CA allows each partner to identify specific third parties it intends to transfer the ownership of its Results to in Attachment (3). Once such parties are identified, other partners waive their right to prior notice and their right to object to such a transfer, which is granted in GA Article 16.4 and Annex 5. However, based on the current version of the CA no partners have identified any third parties to whom they would like to transfer ownership. Therefore, if there is a transfer of ownership to entities that have not been identified in the Attachment 3 of the CA, the provisions of the GA regarding transfer of ownership become applicable. In order to transfer ownership of results generated in the project, the following procedure is applicable:

- i) The partner wishing to transfer ownership must inform the other project partner who have access rights.
- ii) The notification of the transfer must be given at least 45 days in advance. This period is not applicable if it has been agreed otherwise in writing for specifically identified third parties including affiliated entities or unless impossible under the applicable law.
- iii) The notification must include sufficient information on the new owner to enable the partners concerned to assess the effects on their access rights.

- iv) The partners concerned may object within 30 days of receiving notification (or less if agreed in writing); if they can show that, the transfer would adversely affect their access rights.
- v) If a potential transfer would adversely affect access rights, the transfer may not take place until agreement has been reached between the beneficiaries concerned.

7.4.2 Granting of Licenses

The partners are permitted to grant licences to their results (or otherwise give the right to exploit them) to third parties and external entities in accordance with the CA. However, the granting of licenses shall not affect compliance with their obligations. Exclusive licences for results may be granted only if all the other beneficiaries concerned have waived their access rights.

7.5 Steps to Ensure Compliance with the IPR Strategy

In order to monitor the development of results and to ensure that ownership rights are adequately allocated, UNIVE will develop a results ownership document. This document will contain a list of project results and will indicate whether specific results are owned independently or jointly. The document will also include names of the owner(s), the country of the establishment of the owners(s) and information regarding result exploitation.³⁰ The document will be updated regularly based on developments in the project and annexed to the updated versions of the DMP (D8.4, D8.5). In the instance of a joint ownership of results, UNIVIE will draft in collaboration with partners a joint ownership agreement to be signed between joint owners of results. This contract shall outline the terms of the Joint Ownership and ensure compliance with the obligations under the GA and the CA. UNIVIE will monitor the overall compliance with the IPR strategy including the granting of access rights, the protection of pre-existing knowledge and the transfer of rights.

³⁰ [Open Science in Horizon Europe \(europa.eu\)](https://europe.europa.eu)

8. Conclusion

This deliverable outlines the datasets produced and generated in the HOLiFOOD project. The HOLiFOOD lifecycle is described and an analysis of the different stages was performed. It is of utmost importance for the HOLiFOOD consortium to ensure FAIR data management. A general description of how data will be made FAIR in the HOLiFOOD project is provided. Furthermore, section 5 provides detailed information about how partners are handling and securing HOLiFOOD data. Legal and ethical monitoring of the project is of utmost importance. Section 6 highlights some key legal and ethical aspects, when it comes to processing personal data. However, a detailed analysis legal and ethical framework will be provided in D8.2 (M12) and its update (M48). Lastly, this DMP aims to monitor and ensure compliance with the CA and the Intellectual Property Rights (IPR) management strategy that all partners have assented to. This DMP is a living document that will be updated and reviewed regularly in course of the project. Updated Versions will be submitted in M24 and M48.



**holi
food**