MYMATCH

D1.3

Data management plan

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Nature of the deliverable		
R	Document, report	
DEM	Demonstrator, pilot, prototype	
DMP	Data Management Plan	X
OTHER	Software, technical diagram, etc	

Dissemination level		
PU	Public (*fully open*)	X
SEN	Sensitive (*limited under the conditions of the Grant Agreement*)	
EU CI	EU Classified (*eu-restricted, eu-confidential, eu-secret under Decision 2015/444*)	

Deliverable No.	D1.3
Dissemination level	PU
Work Package	WP1
Task	Task 1.3
Lead beneficiary	EQY
Contributing beneficiaries	UCSC
Main authors	Amandine Josseron (EQY), Paola Battilani (UCSC)
Due date of deliverable	31.05.2025 (M6)
Actual submission date	30.05.2025

Quality procedure			
Date	Version	Reviewers	Comments
30.04.2025	V0	UCSC & IBE CNR	Completion of missing information and confirmation of the structure and preliminary content
15.05.2025	V1	External Ethics advisor	Requested additional information on Ethical aspects, addition of privacy notice
30.05.2025	VF	UCSC, IBE CNR & EXUS	Final additions related to Ethics
02.09.2025	VF2	External Ethics advisor	Removed reference to sensitive data collection Mention external ethics advisor Clarify anonymisation/pseudonymisation Nature of open access data Updated consent form

Funded by the European Union

Acknowledgements

This report is part of the deliverables from the project MYMATCH, which has received funding from the European Union's Horizon Europe research and innovation program under grant agreement N° 101181208. More information on the project can be found at www.mymatch-project.eu.

Document's objective and executive summary

The Data Management Plan (DMP) prepared for the MYMATCH project is a living document that will be continuously updated throughout the project. Updated versions will be released by the EQY at Month 24 (November 2026).

In this first version, EQY has outlined the general objectives that all MYMATCH partners should follow for the management of data and metadata, fully aligned with the FAIR data principles and GDPR requirements. As highlighted in the document, since this is only the initial version, more detailed guidelines and methodologies for implementing these principles will be shared with all partners as soon as possible.

The main topics covered in this DMP include:

- A description of the origin, type, and potential re-use of data generated and collected within MYMATCH.
- The process through which data management will adhere to the FAIR principles, ensuring that the data generated in MYMATCH will be Findable, Accessible, Interoperable, and Re-usable.
- ~~Ethical considerations related to data collection and generation, specifying the principles for managing personal and special categories of personal data in compliance with GDPR.~~
- The strategy and tools adopted to guarantee high standards of data processing and security, especially regarding personal ~~and sensitive~~ data.
- An overview of other research outputs produced within MYMATCH.
- An estimation of the resources allocated to ensure that data generation and management comply with the FAIR principles throughout the project.

Table of contents

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

Funded by the European Union

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme ...

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

a mis en forme : Police par défaut, Police :+Corps (Aptos), Vérifier l'orthographe et la grammaire

List of abbreviations

WP – Work Package

DMP - Data Management Plan

EFSA - European Food Safety Authority

ISIMIP – Inter-Sectoral Impact Model Intercomparison Project

FAO – Food and Agriculture Organization

AI – Artificial Intelligence

GDPR – General Data Protection Regulation

FAIR – Findability, Accessibility, Interoperability and Reusability

List of tables

1. Introduction

This document constitutes the **Data Management Plan (DMP)** for the **MYMATCH** project, funded by the European Union (EU) under the Horizon Europe Research and Innovation Programme. The DMP has been developed based on the European Commission's official template, which has been adapted to fit the specific context and needs of the MYMATCH project.

To ensure consistency and harmonization, this document compiles input gathered from all MYMATCH partners through a structured questionnaire answered by the Work Package (WP) leaders with the support of task partners. The responses provided in the data summary section allowed us to gain an overview of the types of data generated and reused within MYMATCH. Contributions to the other sections enabled us to establish an initial framework for data sharing, storage, and processing, as well as to identify existing gaps to be addressed in the following phases.

This is the **first version** of the DMP, representing the early stage of the project's data management. Its main objective is to ensure that all data collected and processed throughout the project lifecycle will comply with the FAIR data principles. The DMP is a living document and will be regularly updated, with one additional version scheduled for release November 2026 (Month 24).

Each partner is individually responsible for ensuring compliance with GDPR (EU Regulation 2016/679) and all relevant regulations related to scientific research. All partners are also responsible for the data they generate within the project. This DMP defines the principles, methods, and standards to be followed throughout the data management process. Each partner, according to their role and contribution, is responsible for complying with all mandatory legislation and the ethical principles guiding the research.

2. Data Summary

2.1 Type and characteristics of generated and reused data

As shown in Table 1, the types of generated data were mainly those which come from questionnaire results, and statistical, modelling and analytical data. Instead, the re-used data are mainly those that come from the literature review. In both cases, the data format will be mainly as documents (word or excel files), images, videos and recordings. The size of the data will be variable, from documents that are sized less than 5MB to those database and tools (e.g., videos for the communications) that will weigh more than 10GB. However, due to the initial stage of the project, for some datasets there are still missing information.

Types of data collected and produced: MYMATCH will primarily utilize both secondary and primary data sources. Secondary data, derived from existing datasets provided by organizations such as EFSA, COPERNICUS, ISIMIP, as well as data gathered from prior EU and national projects, will form the foundation of the project research.

These data will be used in accordance with the licensing and usage terms set by the original data providers. As these datasets are publicly available or institutionally provided, they typically do not involve personal data and thus pose limited ethical concerns. However, appropriate attribution, data integrity, and compliance with reuse conditions will be ensured.

Additionally, primary data collected during the project will supplement analyses: data from workshops and surveys held with stakeholders in **WP3** and **WP10**; from literature in **WP4**; fungi and MY sampling data from field in **WP5**, experimental data from lab in **WP5** & **WP6** and from growth chambers in **WP6**, as well as **weather and climate data** in **WP7**.

In contrast, **primary data** will involve direct interaction with human subjects ~~or the collection of sensitive information~~. In such cases, ethical considerations are paramount. All personal data will be anonymised and handled in full compliance with the General Data Protection Regulation (GDPR – Regulation (EU) 2016/679). Informed consent will be obtained from participants, participation will be voluntary, and data will be processed according to principles of data minimization, confidentiality, and purpose limitation. Where applicable, the project will seek ethical approval from relevant institutional review boards to ensure rigorous ethical oversight.

WP	**Generated data** (description + origin)	**Reused data** (description + origin)
WP1 & WP2	**Description**: Administrative and contact information on members of the consortium and day-to-day project information's such as minutes of meetings.	**Description**: NA
	Origin: Partners	**Origin**: NA
WP3	Survey, interviews and workshops with key stakeholders	NA
	Questions asked to participants in WP3	NA
WP4	NA	We will collect ecophysiological and occurrence data from the scientific literature
	NA	Scientific literature
WP5	• field sampling in natural conditions for fungi and MY monitoring • characterize mycotoxigenic fungi belonging to 3 fungal genera (*Aspergillus*, *Fusarium*, and *Alternaria*) using molecular approaches, including sequencing and chemical profiling, • agrometeorological data related to sampled fields	Data on fungi and MY occurrence in cereals Data on fungi characterisation

	Partners' networks, surveys, research	MYCOKEY project (EC, 2016-2020)
WP6	We will generate datasets containing information about the ecophysiology of fungi. By analysing the data, we will also generate new prediction models in UCSC.	Previous models for some crops will be used to be further improved with the new data. Previous data about the ecophysiology of fungal species will be used.
	Datasets will be generated in different labs across the consortium based on *in vitro* and in phytotrons experiments. Data will contribute to model improvement/development by UCSC	Previous research from UCSC Ecophysiological data from Cranfield does origin from previously research and published papers.
	As WP7 is focused on the Data navigator and predictive models for food system safety, type and characteristics of generated and reused data in WP7 will be fully detailed in a specific section 2.3 below.	
WP8	• AI-driven climate change scenarios and risk assessment data • Generation of missing/duplicated/emulated data based on the data received • Machine learning models trained on historical MY occurrence data. • Decision-support system (DSS) recommendations, including risk levels, mitigation strategies, and early warning alerts.	• Historical climate and MY crop contamination datasets. • Existing agro-climatic and fungal growth models from similar studies.
	• Agroclimatic data from WP7 (climatic, MY and fungi data) • Outputs from WP7's predictive models and scenario simulations. • User requirements and end-users' needs created by Stakeholder (SAB) input based on real-world agricultural and policy perspectives.	• WP7 data navigator and predictive models for food system safety. • MY occurrence data from previous European projects (RAMFIC, MycoKey).
WP9	Training material to familiarise the end users with the project's tools to be demonstrated Data required for the demonstration.	NA
	Material produced by the consortium Data produced by the stakeholder based on their practical needs	NA

WP10	- mitigation guidelines adapted to derived CC scenarios - MY exposure scenarios for different population targets	- possible validation of these updated mitigation measures - policy recommendations to the EU regulatory framework
	WP3, WP4, and WP9	NA
WP11 & 12	• Branding and visual identity data (project's logo, templates, colour palette, and other visual assets) • Target audience analysis data - for reaching different stakeholder groups. • Event data (events organised throughout the project) • Website data (user behaviour, including page views, session duration, and interactions. • Dissemination content (policy recommendations, best practices, and results from campaigns and initiatives) • Social media (engagement metrics - likes, shares, comments), follower growth, and to be used to evaluate the reach and effectiveness of the project's online presence • Collaboration and synergy (data related to the establishment of synergies with other projects and organisations - records of joint activities, agreements, and collaborative efforts). • Feedback and evaluation (feedback will be collected from participants, stakeholders, and audiences - survey responses, evaluation forms, and direct feedback from events) • Creative outputs (related to the creation of artistic outputs and the documentation of workshops and creative activities, records of the creative process, participant involvement, and the final products produced).	• External events data (events organised by partners and other projects) • Content created by partners (i.e. pictures)

Table 1: Type and characteristics of generated and reused data

2.2 Purpose and utility outside the project of the generated and reused data

The purpose of data generation and re-use together with the utility of data outside the project are shown in Tables 2 and 3. In the MYMATCH project, data is generated to support project coordination, stakeholder engagement, and the development of predictive models and an AI-driven risk assessment platform. The collected data informs the understanding of stakeholder needs, monitoring of fungal populations and contamination risks, and enhances model accuracy under evolving environmental conditions, while existing data is reused to establish baselines, validate models, and ensure alignment with regulatory frameworks. Externally, this data proves valuable to public institutions, policymakers, food industries, farmers, consumers, and even climate scientists and AI communities by aiding in the development of prevention strategies, risk management tools, and decision-support systems for agricultural safety and supply chain optimization.

WP	What is the purpose of the data generation	What is the purpose of the data reuse?
WP1 & 2	Contact project members and follow up on actions of the project.	NA
WP3	Collecting insights to : • Best orient project results and exploitation • Understanding levels of interests and acceptability of stakeholders towards the project results • Getting insights to orient the platform development	NA
WP4	NA	Occurrence data will be used as the state of the art for the baseline contamination scenario. Ecophisiological data from the literature will be used to generate new models or for testing the predictive models generated within the consortium

WP5	• addressing knowledge gaps in mycotoxigenic fungi and related MYs prevalence in Europe • seeking to monitor the spread of highly toxigenic producer strains across seven studied countries • identifying genetic traits that can track the composition and dynamics of key fungal populations • comparing fungal occurrence with data collected in past years/projects	• create an historical dataset on fungi and MY levels in cereals in Europe • Compare data with similar previously collected information, to understand fungal shift due to CC • compare past and current fungi community and their metabolites to give support for model validation and demonstration.
WP6	~~The purpose of the data generation is to p~~Produce the necessary knowledge to produce an informed development of the prediction models, so they can produce good predictions under the new forecasted environmental conditions and extreme events.	Similar to the previous, we will use previous prediction models to develop further their performances to produce good results with the new environmental conditions
WP8	• To develop an AI-driven platform that will provide its users with a risk assessment for MY contamination for Improved Preparedness. • To support decision-making by generating predictive insights for farmers, policymakers, and food industry stakeholders with the use of AI. • To ensure stakeholder-driven scenario building, incorporating SAB feedback to improve risk prediction and mitigation strategies	• To enhance the accuracy of the created AI models • To validate predictive models against historical data and stakeholder insights. • To align MYMATCH outputs with policy needs, and agricultural risk assessments

WP9	The training materials are developed to facilitate the effective uptake of the MYMATCH platform by diverse end-user groups, ensuring they are able to understand and apply the tools in real-world settings. These materials aim to bridge the gap between technical outputs and user-friendly application. The demonstration data, including stakeholder-generated data during the validation and piloting phases, serve to adapt the platform to user-specific needs and validate its usability and performance under realistic field conditions.	NA
WP10	• Ensure stakeholders engagement and carry out the collaborative activities with the FS4EU platform and other national and EU projects and initiatives • transmission of results to public institutions in strict collaboration with EFSA	• involvement of stakeholders, by consultation or contribution in the design, feasibility, and possible validation of MYMATCH achievements • to improve the preparedness-level of the EU regulatory framework
WP11 & 12	~~The purpose is to m~~Measure and enhance the project's communication efforts, ensuring that its findings and results are effectively disseminated to the right audiences and that stakeholder engagement is continuously monitored and improved. Data in WP11 & 12 includes a wide range of materials and metrics designed to measure the effectiveness of the project's outreach and engagement strategies.	

Table 2– Purpose of the generated and reused data

WP	To whom might your data be useful ('data utility') outside the project
WP1 & 2	None. Information for internal use only.
WP3	Public Institutions (i.e. EFSA, FAO, National Agencies) and relevant stakeholders (i.e. consumers, farmers and industries)

WP 4	None. Internally used only. Data will be used for the preparation of systematic reviews.
WP5	• Developing prevention strategies, including the use of predictive models to assess MY occurrence risks in different scenarios • To increase awareness on food safety aspects; • To explore strategies for adapting food safety systems to climate change impacts; • To improve networking and knowledge exchange on food safety.
WP 6	The generated data could be useful to develop prediction models. These models will be further developed and tested in WP7.
WP8	The main target group for whom the data are targeted are those that have been documented in the GA and are part of the Stakeholder Advisory Board (SAB). We have envisioned the use from: • food industries, • farmers, • policymakers, • consumers Furthermore, additional groups that may benefit from the data gathered and created from the project implementation are: • **Climate Scientists** that will analyze climate change's impact on MY occurrence in food crops that are considered in the project (wheat, maize, and tomato) or other relevant ones. • **AI & Data Science Communities**: Improving AI-based decision-support tools for agricultural risk management. **Supply Chain Managers**: Optimizing their production strategies based on MY risk predictions and modify according to their supply chains to ensure a smooth provision of goods.
WP9	Training materials can support wider dissemination and replication of the MYMATCH platform beyond the project's immediate stakeholder group, particularly in similar agro-climatic or regulatory contexts. Demonstration and validation datasets may be valuable to policy-makers, national food safety authorities, and agricultural advisory bodies to inform risk assessment protocols, regional adaptation strategies, and training curricula for digital agriculture and food safety management. These resources could also serve as reference material in academic and vocational training programs focused on agro-digital transition, decision support systems, and AI applications in agriculture.
WP 10	Public Institutions (i.e. EFSA, FAO, National Agencies) and relevant stakeholders (i.e. consumers, farmers and industries)
WP1 1 & 12	Stakeholders willing to disseminate the same type of information as the ones generated by the communication team of the MYMATCH project.

Table 3– Data utility outside the project

2.3 Focus on WP7: Data navigator and predictive models for food system safety

2.3.1 Section 1: Data from weather stations

This section includes observational meteorological data from national, regional, and local weather stations. Data sources range from public agrometeorological networks to private or partner-managed weather stations. These datasets provide ground-truth reference for climate indicators and are used for model calibration, downscaling, and validation.

Reused:

- National and regional agrometeorological services (e.g., https://www.aemet.es/en/portada , https://www.ipma.pt/en/agrometeorologia/mapas , https://www.metoffice.gov.uk/, https://meteohub.mistralportal.it/app/datasets https://www.met.hu/idojaras/agrometeorologia/ https://www.hidmet.gov.rs/eng/meteorologija/agro.php https://www.nibio.no/en/services/lmt)
- International open datasets such as ECA&D [https://www.ecad.eu]
- HadISD [https://www.metoffice.gov.uk/hadobs/hadisd/]
- ASOS [https://mesonet.agron.iastate.edu/ASOS/]
- GHCN-D [https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily]
- JRC Gridded Agro-Meteorological Data in Europe – 25x25 km resolution [https://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx]

Generated:

- Private or partner-owned weather stations (currently under identification)

Purpose of generated data:

Created to extend spatial coverage and ensure high-resolution observations where official datasets are sparse, especially from private or partner-managed stations.

External usefulness:

Can benefit regional meteorological services, ag-tech developers, and local advisory systems interested in localized weather analytics or model refinement.

2.3.2 Section 2: Gridded datasets

Gridded datasets provide spatially continuous representations of climate variables derived from interpolated station data, satellites, or reanalyses. They are useful for regional-scale modeling, filling spatial gaps, and performing statistical analyses.

Reused:

- E-OBS [https://www.ecad.eu/download/ensembles/ensembles.php]
- CRU TS 4.01 [https://crudata.uea.ac.uk/cru/data/hrg/]
- CHIRPS [https://www.chc.ucsb.edu/data/chirps]
- AgMERRA and AgCFSR [https://data.giss.nasa.gov/impacts/agmipcf/]
- JRC MARS (Agri4Cast) – 25x25 km grid [https://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx]

2.3.3 Section 3: Satellite datasets

Satellite datasets offer continuous spatial and temporal data on surface variables like solar radiation, sunshine duration, and temperature. They complement ground-based data and improve input resolution for models.

Reused:

- CM SAF SARAH-2 [https://wui.cmsaf.eu/safira/action/viewProduktSearch]

- HelioCLIM-1/3 [http://www.soda-pro.com/web-services#radiation]

2.3.4 Section 4: Reanalyses

Description:

Reanalysis datasets combine observations with model output to generate consistent climate datasets across space and time. They are used for retrospective analysis, model validation, and trend analysis.

Reused:

- ERA5, ERA5-land, Agrometeorological indicators [https://cds.climate.copernicus.eu/]
- NCEP2 [https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.html]

Generated :

- High Spatial Resolution Reanalyses Data

Purpose of generated data:

High-resolution reanalysis fields are developed to improve spatial accuracy for local applications, especially in agricultural risk assessments.

External usefulness:

Applicable for insurance models, adaptation policy simulations, and environmental monitoring organizations requiring localized long-term climate reconstructions.

2.3.5 Section 5: Climate models

This section includes outputs from global and regional climate model simulations used to assess historical and future climate conditions. Models include historical runs, scenario projections, and downscaled outputs.

Reused:

- EURO-CORDEX [http://www.euro-cordex.net/]
- Med-CORDEX [https://www.medcordex.eu/]
- CMIP6 – ISIMIP Portal [https://www.isimip.org/]
- UERRA [https://uerra.eu/outreach/data-and-products.html , https://cds.climate.copernicus.eu/]

Generated :

- High Spatial Resolution Climate Data

Purpose of generated data:

Custom downscaled and bias-adjusted climate datasets tailored for site-specific agro-environmental simulations.

External usefulness:

Highly valuable for climate service providers, adaptation planners, and other modeling consortia seeking high-resolution projections over agricultural regions.

2.3.6 Section 6: Seasonal forecast

Description:

Ensemble-based seasonal forecast datasets from major meteorological agencies, used for predictive modelling and adaptation strategies on seasonal scales.

Reused:

- C3S Multi-model seasonal forecast [https://climate.copernicus.eu/seasonal-forecasts]

Generated:

- High Spatial Resolution Climate Data

Purpose of generated data:

Customized forecast maps and indicators for site-specific agro-climatic insights and modeling.

External usefulness:

Potentially useful for regional early warning systems, precision farming platforms, and supply chain risk analysts.

2.3.7 Section 7: Mycotoxins and fungi data (from WP4, WP5, WP6)

Includes observational, experimental, and laboratory data related to fungal species ecology, co-occurrence, and mycotoxin concentrations (e.g. Aflatoxins, Fumonisins, Alternaria toxins). These are essential for WP7 predictive model improvement/development.

Reused and generated:

- Collected and processed by WP4 (harmonized literature data), WP5 (field conditions data), and WP6 (experimental trials in controlled environment).
- Data will be integrated in WP7 via the centralized AgroSat platform database.

Purpose of generated data:

Includes unique datasets from experimental trials and field campaigns used to improve fungal risk assessment tools.

External usefulness:

Valuable to food safety agencies, crop breeders, and risk modelers interested in integrating biotic stress factors into yield and safety predictions.

2.3.8 Section 8: Suitability analysis and crop model outputs

Outputs from crop growth and suitability models including agro-climatic zones, crop distribution, and environmental suitability under current and future conditions. Includes data for phenology, crop calendars, irrigation mapping, and model parameters. Crop suitability assessments will also exploit crop models implemented and validated by project partners, ensuring consistency with local agronomic conditions and enhancing the accuracy of spatial predictions.

Reused:

- Yearly modeled crop area in EU – [https://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx]
- European Irrigation Map (EIM2010) – [https://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx]
- Gridded winter soft wheat phenological database – [https://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx]
- EU Crop Calendar (Winter wheat, maize, rice) – [https://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx]

Generated:
Crop suitability maps and agro-climatic zoning layers

Purpose of generated data:

To assess how climate change alters the suitability of specific crops across regions

External usefulness:

Useful for agro-advisory services, regional planners, land-use strategists, and sustainability certification schemes needing scientifically validated agronomic predictions.

3. Fair data

In this section, the procedure for managing data in accordance with the FAIR principles is outlined. Before detailing these principles for the MYMATCH project, an overview of the Open Science practices that the Consortium will implement is provided.

Open methodology

The project will adopt an open methodology, meaning that whenever possible the specific approaches used in MYMATCH will be described, discussed, and shared externally. This openness will enable external expert panels—such as the European Food Safety Authority —to provide valuable feedback and raise awareness about the project. Additionally, strategies for stakeholder engagement will be developed through direct discussions during user co-creation workshops. All implemented methodologies will be thoroughly documented and made available in dedicated open access publications.

a mis en forme : Non Surlignage

Open Access to Research Outputs

Research outputs, including data analysis, new methodologies, and scientific findings, will be accessible and shared throughout the project, with the exception of the code used in AI algorithms. Raw data and results, once anonymized and compliant with privacy requirements, will be archived on the MYMATCH AI MY platform in alignment with the FAIR principles to ensure maximum discoverability.

Moreover, scientific articles, conference presentations, reports, and potential white papers will be published as open access to support both applied and fundamental research, as well as to inform future policy recommendations—thus maximizing the project's overall impact. ~~Prior to releasing any project data, the Consortium will ensure full compliance with relevant legislation, including GDPR and Regulation (EU) 2016/679~~.

MYMATCH will share knowledge and tools as early and widely as possible in an open and cooperative way with FAIR and open by default research data, and engage citizens, civil society, and end-users, where appropriate. To ensure responsible management of research outputs, the partners will adopt the 'gold' open access model, making scientific publications immediately accessible online free of charge or by having each affiliated partner cover the relevant cost. In cases where the 'gold' access model is not applicable, the 'green' model will be used by supplementary publishing of relevant articles to an online repository in consultation with the publisher, in case an embargo period is needed.

Finally, the collaboration with EFSA will foster open-science practices and knowledge sharing.

3.1 Making data findable, including provisions for metadata

A common data management platform (**T7.1**) and a SharePoint for "light data" will be used to store and back up data collected/generated all along MYMATCH. The dataset will be accompanied by comprehensive metadata aimed at enhancing its discoverability, citation, and reusability. Rich metadata will offer valuable contextual information crucial for interpreting the data, while also facilitating automated analysis by machines. The data will be stored in servers physically hosted in EU countries, to make sure the EU laws apply. Furthermore, data that is not submitted to IPR, as well as outcomes of the project, will mainly be exploited in publications (each with its own DOI), in line with open science practices. Relevant data will be also available on the project website and social media, as well as stored on a trusted repository (e.g. Zenodo, European open science cloud), in a standard format abiding by the open research data requirements.

3.2 Making data accessible

Using the secure project SharePoint and data platform, partners will have full access to relevant collected/generated data. Upon legitimate request, partners may grant specific access to third parties with respective partners' approval. ~~The access to sensitive raw data collected from stakeholders, interviews and similar primary data generated within the project will be limited to the partners in charge.~~

3.3 Making data interoperable

Open data standards will be used when possible, including the ones that will be used for captured data. The produced data will use community agreed schemas, controlled vocabularies, keywords, thesauri or ontologies where possible in order to be interoperable and be integrated with other data, applications and workflows. Moreover, the data produced during the project will be discussed with EFSA to ensure data sources and data collection and/or data generation are in line with prioritized emerging issues. As well, the predictive models will be developed with the view of being reused by EFSA.

3.4 Increase data re-use

Partners will deposit technical and other professional information in the European Open Science Cloud (EOSC) to enable storage, sharing, processing, analysis, and re-use of research outputs (data and software) across borders and disciplines. The usage of EOSC and other trusted repositories (e.g., Zenodo) will enhance the reproducibility of MYMATCH research outputs, limit waste of resources, increase the reusability and reliability of research, and thus strengthen European R&I. MYMATCH's open science strategy will promote innovation and transfer through reusable and reproducible results, enable increased quality and efficiency of research, create more trust in science, and lead to important breakthroughs by European researchers, boosting knowledge and competitiveness in Europe.

The data generated will be meticulously documented, accompanied by transparent licensing and provenance details. To facilitate proper interpretation and reanalysis by other researchers, a README file will be included with the dataset. In terms of licensing, the data will adhere to clear licensing terms governing their reuse, such as Public Domain, Attribution, Non-commercial, No Derivatives, or others. As an example, MYMATCH intends to utilize Creative Commons licenses, specifically the Attribution (CC BY) license and Creative Commons Zero (CC0) license.

Strategy for data re-use will be further defined in collaboration with the updates of the communication, dissemination and exploitation plans (first draft produced at M6- May 2025 and updates at M18 – May 2026, M36 – November 2027 and M48 – November 2028).

4. Data Protection, Storage, and Rights Management

Aligned with the data governance principles previously outlined, this section describes the measures implemented to ensure lawful, secure, and FAIR-compliant data management.

Agrosat is a platform designed with both technical and organizational safeguards to provide intrinsic security for data and data communications. Additionally, specific mechanisms have been developed to meet the requirements of the General Data Protection Regulation (GDPR) and the FAIR data principles.

4.1 Data Minimisation and anonymisation~~Anonymisation~~

Agrosat enforces strict data minimisation policies, collecting only the data necessary to deliver its services. Access to advanced features requires user registration, including the provision of identifiers such as email and username. When a user deletes their account, Agrosat applies ~~pseudonymisation~~ anonymisation by replacing personal identifiers with a system-generated user ID. Historical data are retained for analytical and research purposes, but are no longer traceable to the original user.

4.2 Access Control and Interoperability

Agrosat integrates data from diverse external sources accessed via specific API calls that use authenticated and encrypted communication protocols. This ensures secure connectivity with third-party platforms and datasets.

Many Agrosat data products are generated in real time and are not stored persistently. More complex and data-intensive products are pre-processed and stored in the platform's database to improve performance and accessibility.

Access control is managed through a role-based system that assigns permissions according to user roles (e.g., users, decision-makers, applications, platforms). Authentication follows a dual-token model, aligned with OAuth 2.0 specifications:

- Access Token: used to make API requests.
- Refresh Token: used to regenerate the Access Token upon expiration.

This approach enables secure, continuous interaction with the platform. Agrosat also supports interoperability by exposing WebService APIs that allow both data retrieval and submission, ensuring smooth integration with other systems and services within the MYMATCH ecosystem.

4.3 Data Sharing and Long-Term Protection

User-related data are stored in the Agrosat Geo-Database, along with ancillary and foundational datasets, including Copernicus ERA5 and Sentinel-2.

This Geo-Database, along with the WebService APIs and predictive models, is hosted on a network of specialized Virtual Machines (VMs) managed within CNR-IBE's primary IT infrastructure in Florence (Italy).

To ensure long-term data and service protection:

- Backups of the VMs are performed every 4 hours.
- Replicated backups are created every two days and stored at IBE's secondary site in Bologna (Italy).
- In the event of a disaster, full VM recovery can be achieved within 60 minutes to 5 working days, depending on the severity of the incident.

Within the MYMATCH project, data sharing is managed via user accounts and role-based access. External users can access available services and resources through their registered accounts.

5. Other research outputs

In addition to the management of data, a plan is defined below for the management of other research outputs following the FAIR principles and strategy defined above.

- **New fundamental knowledge on mycotoxigenic fungi prevalence in Europe in cereals and tomatoes**: publications will be made available through golden open access (OA) and invited presentations and conferences.
- **Set of predictive models for predicting MY occurrence**: They will be widely shared among their end-users, through the project website, presented in journals and in conferences. After validation, models will be available for free download (use and code source) with full documentation.
- **Climate change scenario builder**: Climate Change scenario will be widely presented in peer-reviewed journals, conferences, shared with EFSA and JRC and available in OA.
- **MYMATCH AI MY platform:** The platform will be integrated into the EXUS Financial Suite (EFS). It will allow EXUS to commercialise the upgraded product via its sales

network (more than 35 countries worldwide). EXUS and partners involved in the development will elaborate an exploitation agreement with EFSA under fair and reasonable conditions and will co-own according to the IP rights, defined in the context of the IPR Management (WP1&2). Thus, EXUS will lead the post-project exploitation activities, but all involved partners will be able to claim royalties from sales. It could ultimately lead to the opening of a new market and product category for EXUS, strengthening platform use and maintenance, and its continuous enrichment.

- **Guidelines for practitioners for mitigation measures adapted to Climate Change**: Widely shared in OA, on the project website, social media, and in relevant magazines and channels consulted by farmers.
- **Policy recommendations**: Shared by each partner in OA to the relevant public bodies of their country, as well as at EU level. They will also be presented to policymakers and publics institutions as EFSA and FAO.

6. Allocation of resources

The resources for making the DMP, updating it throughout the project and monitor the data management are foreseen in the project budget are allocated to **EQY** and **UCSC** that is responsible for the data management in the project. Furthermore, during the project, **UCSC** as coordinator has~~will~~ appointed a dedicated ethics advisor to provide insight on ethical issues related to data management of personal data and their possible use in AI algorithms. The personal data will be dealt in accordance with ethical principles and applicable international, EU and national laws, in particular EU Reg. 679/2016to limit the impact on the persons concerned and ensure data quality and confidentiality. The research is compliant to the GDPR 2016/79~~The GDPR 2016/79 is taken into consideration and compliance under its articles will be ensured~~. All necessary measures to achieve compliance will be undertaken (e.g. consent forms, data protection officer appointment, etc.). After the end of the project, costs for long-term data preservation will be undertaken by each partner for the data backed-up on their servers. Furthermore, costs arising from preservation of the project outcomes on the project website will be undertaken by **EQY**.

A Data Protection Officer has been appointed by UCSC in the person of Avv. Zanatti. She will review the Grant Agreement to prevent any data protection issue and support the consortium in managing any data protection issue during the lifetime of the project.

7. Ethical data processing

The project appointed an Ethics advisor to provide insight on ethical issues related to data management of personal data.

~~6.~~

a mis en forme : Police :(Par défaut) Montserrat, 11 pt, Couleur de police : Automatique

a mis en forme : Police :(Par défaut) Montserrat, 11 pt, Couleur de police : Automatique, Anglais

a mis en forme : Italien (Italie)

a mis en forme : Normal

~~6.2~~7.1 Non EU countries

Some data will be transferred to non-EU countries through the partners Norwegian Veterinary Institute(NVI, located in Norway), the Maize Research Institute (MRI, located in Serbia) and Cranfield University (CRANFIELD, located in the UK). MRI (Serbia), NVI (Norway), and CRANFIELD (England) will take part all over the course of the project, in particular for sampling, fungi identification and characterization, and experiments. They will also be involved in the demonstration of the models and the platform. The partners have confirmed that the ethical standards and guidelines for Horizon Europe will be rigorously applied, regardless of the country in which the research is carried out. In case activities undertaken in non-EU countries raise ethics issues, the consortium will gather and ensure that the EU rules are respected.

~~6.3~~7.2 Use of artificial intelligence for personal data

No individual-level profiling is performed or intended. The AI-based risk prediction and decision support tools are designed to operate at a **crop-system level**, and their outputs relate to **environmental, climatic, and agronomic risk factors**, not to data subjects or personal behavioral patterns.

While the tool aims to offer **tailored recommendations**, this tailoring is based on **agricultural context (e.g., crop type, geography, climate scenario)** rather than user-specific profiles. The information input from stakeholders (via questionnaires or workshops) is aggregated and used to **inform general system requirements** and usability—not to train models that adapt to or infer individual user behaviour.

In that:

- The **effectiveness of the system does not rely on profiling** of individuals, but on **modeling the relationship between environmental variables and fungal/mycotoxin dynamics**.
- The **system cannot make decisions 'on' individuals**, nor does it process personal attributes to generate outputs.
- All data collected from human participants will be processed in line with **GDPR**, with ~~**pseudonymisation**~~ **anonymisation, and minimisation** applied.
- No automated decisions are made about participants; **recommendations are general, context-based, and advisory**, with **human oversight (human on the loop)** embedded in all use cases.
- Regarding **dual-use concerns**, the AI tools do not provide any functionality that could reasonably be used to control or target individuals, manipulate personal behaviors, or be repurposed for surveillance or security applications.

~~6.4~~7.3 Collection of ~~sensitive~~ data through interviews and questionnaires

~~6.4.1~~7.3.1 Structure and Thematic Areas

The interviews/questionnaires are structured into five thematic sections:

- Climate Change Perceptions and Challenges – explores farmers' awareness, concern levels, and experience with climate-related events impacting agriculture.
- Mycotoxin Contaminant Views & Food System Awareness – assesses knowledge, monitoring practices, and challenges related to mycotoxin contamination.
- Digital Knowledge – evaluates familiarity with and use of digital tools in agricultural practices.
- Willingness to Engage in Multi-Actor Platforms – explores interest in collaborative digital systems for food safety and risk management.
- Demographic Information – includes broad, non-identifying categories (e.g. age range, education level, farm size, production type, region).

~~6.4.2~~7.3.2 Nature of Collected Data

All data is anonymous by design: no names, contact details, or specific location coordinates are collected. Demographic data is collected in generalized categories only (e.g. "region of operation" rather than town; age as number, not date of birth). The content focuses on perceptions, awareness, practices, and willingness to engage—not on private or personally sensitive details. Questions that could involve indirect identifiers (e.g. region + farm size) will be aggregated and anonymized during analysis to prevent re-identification.

~~6.4.3~~7.3.3 Data Anonymization Measures

All responses are anonymous at the point of collection.
Before analysis, data is reviewed to ensure removal of any potentially identifying combinations.
Results will be reported in aggregate only, e.g., by country or region, not by individual respondent.

~~6.5~~7.4 Recap of data classification in MYMATCH

Classification Dimension	Status in This Project
Primary Data	Yes – collected during research activities as well as via interviews and questionnaires
Secondary Data	Yes – from publicly available literature and datasets
Personal Data	Not collected – all data is anonymous and generalized
Anonymized Data	Yes – no directly or indirectly identifiable information is retained

Restricted or Classified Data	Not applicable – no data under EU restriction or classification
Open Access Policy	Followed – data and publications will be made openly available where possible
Copyright / Intellectual Property	Governed by the Consortium Agreement; proper licensing and citation practices apply
Dual-Use Potential	None – no outputs with application risks
Research Integrity & Authorship	Ensured – proper attribution, transparency, and ethical standards maintained

Table 4: Recap of data classification 1

~~7.~~8. Conclusion and next steps

This first version of the MYMATCH project Data Management Plan (DMP) aimed to set up the initial guidelines to manage data and metadata that the project will produce (both using generated and re-used data), making them compatible with the FAIR principles. The DMP will be a continuously updated document.

The initial exchanges with partners, including survey results, helped us understand which topics should be further explored and which still need clarification. For this reason, specific meetings will be organized to take common decisions that will be approved by the Coordination team, before being validated by the Consortium.

~~8.~~9. Annex:

~~8.1~~9.1 Summary of WP7 Online Seminar on Data Harmonization

On May 7, 2025, the WP7 data seminar was held online with 23 participants representing all involved work packages. The objective was to align on key aspects of data harmonization, integration, and platform support within the MYMATCH project.

WP4 provided a comprehensive overview of their data workflows, with a clear definition of methods, data formats, and metadata standards. These elements were discussed and formally agreed upon between WP4 and WP7 partners and will serve as a structural reference for data organization across other thematic areas.

DB_origin	Ana

sampCountryorigin	RO
sampContinentorigin	EUROPE
sampContinent	EUROPE
paramType	DON
paramTypeO	DON
sampCountry	RO
sampRegion	Arad;Timis
sampInfo_latitude	NaN
sampInfo_longitude	NaN
sampInfo_altitude	NaN
sampYear	2010
sampYearIncreas	0
sampSize	53
sampMethod	Comission Regulation (EC) No. 576/2006
sampPoint	preharvest
sampPointInfo	farm
sampMatType	feed
sampMatbased	wheat
sampMatbasedO	wheat
sampMatInfo	grains
sampMatCode	NaN
growingSystem	NaN
anMethRefId	ELISA
resUnit	µg/kg
resLOD	25,69
resLODinfo	NaN
resLOQ	44,6
resLOQinfo	NaN
exprResPerc	NaN
ToTresValUncertSD	NaN
meanTot	NaN
POSresValUncertSD	NaN
meanPos	2263,1
median	NaN
medianinfo	NaN
min	294
max	3390
MAJ_LOD	73,08
IQRmin	NaN
IQRmax	NaN
Concentration	NaN
DataSampLevel	0
Co_occurrence	1
Ref	Alexa, E. et al, 2013
Ref_unique	Alexa, E. et al, 2013

Funded by the European Union

During the WP5 session, the planned activities and data streams were clarified. Importantly, the adoption of a similar scheme to WP4 for field-level mycotoxin data was discussed, aiming to streamline integration and standardization. A dedicated dashboard for field data management will be made available on the project platform and will be adaptable to the specific needs of each partner. As for weather data, the shared formats discussed during the meeting can be refined and adjusted to fit partner-specific data pipelines.

MYMATCH

Data Harmonization Seminar

MYMATCH WP5 T5.1

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1

MYMATCH

MYMATCH

T5.1: Sampling in demonstrator fields and MYs detection – Leader: MRI

Data formats that will be used :

1. Field information
2. Agronomic practices data
3. Meteorological data
4. Fungal occurrence in crops
5. Mycotoxins occurrence in crops

MYMATCH

1. Field information

N°	crop	Locality (ID)	GPS coordinates	Field area (m^2)	N° of samples in the field (folowing the model X or W)	Approx.distance from the meteo station (km)	Sampling year
1.	maize				20		1st (2025)
2.							
3.	maize				20		2nd(2026)
4.							
	wheat				50		1st (2025)
	wheat				50		2nd(2026)
	tomato				50		1st (2025)
	tomato				50		2nd(2026)

Funded by the European Union

MYMATCH

MYMATCH

2. Agronomic practices data

Locality ..... **Year** ..... **Sample N°**.....
Hybrid..... F.A.O class..... Brand.....

Soil type.....(sand %.....; clay%.....)

Preceding crop.....

Plants residue on soil □ yes □ no **Tillage** □ **No tillage** □

Ploughing □ winter □ spring □ summer □ autumn Depth (cm) □ 25 □ 30 □ 35 □ 40 □ 45 □ other:.....

Seeding date..... N°plants/m^2 □ 5.5 □ 6.0 □ 6.5 □ 7.0 □ 7.5 □ other.....

Silk emergence (date).....

Harvest date..... **grain moisture at harvest**.....

		light <10%	medium 10-30%	severe >30%
Damages	insects.....			
	hail.....			
	lodging.....			

4 **20/05/2025**

MYMATCH

2. Agronomic practices data

Manuring

Commercial product & N, P_2O_5 and K_2O content	Dosage (Kg/ha)
.....	
.....	
.....	
.....	

Weeds control

Commercial product .....
Active ingredient.....
Date.....

Irrigation

Type . □ overhead
□ furrow

N° ..... Date.....
.....
Amount of water/ each irrigation mm).....
.....

Insects control

□ yes □ no

Commercial product..... Active ingredient.....
Date .....

5 **20/05/2025**

Funded by the European Union

MYMATCH

3. Meteorological data

Daily format

Date	AirT_max (°C)	AirT_avg (°C)	AirT_min (°C)	Rh_max (%)	Rh_avg (%)	Rh_min (%)	Rain_daily cum (mm)	Rain_hourly _max(mm)	SoilT_ma x (°C)	SoilT_avg (°C)	SoilT_min (°C)	W_speed_ avg (m/s)	W_speed_ gust (m/s)	W_dir_prevaili ng(°)	Solar radiati on_av g (w/m 2)	Leaf wetness (total wet hours)
YYYY/MM/DD	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xxx.x	xxx.x	xx

MYMATCH

5. Mycotoxins occurrence in crops

Locality..... Region.....

Crop (maize/ wheat/tomato)	AFLA	FUM	DON	T-2	HT-2	TeA	ALT	AOH	ATX	TEN	AME
Number of tested samples											
Positive (%)											
Average MY concentration of positive samples											
Average MY concentration of total samples											
Range of MY conc											

Funded by the European Union

MYMATCH

MYMATCH

3. Meteorological data

Hourly format

Date	Time (UTC+0)	AirT (°C)	Rh (%)	Rain (mm)	SoilT (°C)	W_speed	W_dir(°)	Solar radiation (w/m2)	Leaf wetness (%)
YYYY/MM/	HH	xx.x	xx.x	xx.x	xx.x	xx.x	xxx.x	xxx.x	xx.x

7 20/05/2025

MYMATCH WP5 T5.2

MYMATCH

T5.2 – data to be generated and shared

T5.2: Identification and characterization of toxigenic fungal strains isolated from cereals and tomato – Leader: NÉBIH – Participants: MRI, CNR, UNIPR, UMINHO, IFC – M6-M36

(i) Representative strains from the experimental fields will be deposited in culture collections at CNR, NÉBIH, and UMINHO.
(ii) CNR, IFC, MRI, and UMINHO will sequence DNA markers for the taxonomical identification of fungi. Analysis will be set up by CNR, NÉBIH, IFC, and UMINHO.
(iii) They will contribute to the aim separately and partially to reach final objective as consortium, with specific contributes: IFC will be in charge of fungal species identifications and characterizations; NÉBIH will focus on fungal whole genome sequencing, genome assembly, annotation, phylogenomics and stress gene identifications; setting up and maintaining *A. flavus* strain collection will be assigned to IFC; UMINHO will contribute with some genome sequences and fungal species identifications; CNR will contribute with fungal species identifications, comparative genomics, secondary metabolites gene annotation.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1

MYMATCH

MYMATCH WP5 T5.2

MYMATCH

(iv) Data will converge to identify genetic and metabolic modifications possibly related to adaptation to climate change. This will be obtained comparing selected strains against representative strains isolated over at least a 10-year time frame and available in fungal collection at CNR. In addition, Polyphasic and/or –omics approaches will be used in the mentioned strains (selected from field sampling in MYMATCH and those from fungal collections) to verify fungal response to climate change stress, focusing on genetic traits related to chemotyping and MYs production.

(v) Knockout of stress related genes in a key species will be conducted to verify their role in the same extreme environmental conditions tested in WP6. The results will be combined with the field sampling in a MY occurrence report (D5.1, M36) (part of PR1).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

MYMATCH WP5 T5.2

MYMATCH

Agreed

- partner will **ship** strains at least to 1 colture collection, according to Nagoya protocol.
- DNA markers to sequence:

Aspergillus	Tef, CaM, ITS
Fusarium	Tef, CaM, ITS
Alternaria	GPD, Alta1, ITS

	N° of strains
NFCSO	200
CNR-ISPA	200
UMINHO	Up to 200
MRI	**Up to 200**
Total	Around 800

CNR, IFC, MRI, and UMINHO will sequence DNA markers for the taxonomical identification of fungi. Analysis will be set up by CNR, NÉBIH, IFC, and UMINHO.

KPI2: B= At the end of MYMATCH, at least 4 genetic traits. A= As multiple genetic traits can be identified by genome sequence analysis, additional traits can be retrieved from genomes, at a rate of 1 every 2 years.

primer names 5'-3'	reference	Aspergillus	Fusarium	Alternaria
tef1		x	x	x
tub2		x	x	
cmd5		x		
tub2			x	
rpb2-7.1R				x
rpb2-6.4				x

Type of data:
- isolates identified at sp. level *via* morphological features (around 800) and at sp. level *via* molecular markers (around 400)

Data sources:
- isolates from field sampling, monospore cultures, visual and microscopic observations, PCR amplicons and DNA sequencing

Format:
- XLS files, containing isolate IDs, species names, DNA sequences (FASTQ), GenBank Accession Nos.

Metadata:
- geographic origin of isolate, date of collection, host, strain collection IDs

Sharing:
- through shared cloud space, GenBank

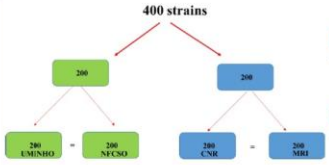
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

MYMATCH WP5 T5.2

MYMATCH

Culture Collections Involved

- CNR (Italy)
- NFCSO (Hungary)
- UMINHO (Portugal)
- MRI (Serbia)



Representative strains from the experimental fields will be deposited in culture collections at CNR, NÉBIH, and UMINHO.

KPI1: B= At the end of MYMATCH, 300 representative strains will be collected totally for the 3 crops. A= As same species can occur on same crop and different countries, multiple representative strains for single species will be collected if detected (rate 20/year).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

4

Type of data:
- selected and characterized representative strains (400) will be deposited in at least two culture collections within the consortium (UMINHO-NFCSO/NÉBIH, CSR-MRI)

Data sources:
- representative strains for single species will be deposited after consortial decision (20/species/year)

Format:
- XLS files, containing WDCM (World Data Centre of Microorganisms) IDs, authorization to issue strains

Metadata:
- geographic origin of isolate, date of collection, host, strain collection IDs

Sharing:
- without any limitations within the consortium, and with MTAs (Material Transfer Agreements) on request outside consortium, compliance with and enforcement of Nagoya protocols

MYMATCH WP5 T5.2

MYMATCH

Data will converge to identify genetic and metabolic modifications possibly related to adaptation to climate change. This will be obtained comparing selected strains against representative strains isolated over at least a 10-year time frame and available in fungal collection at CNR. In addition, Polyphasic and/or –omics approaches will be used in the mentioned strains (selected from field sampling in MYMATCH and those from fungal collections) to verify fungal response to climate change stress, focusing on genetic traits related to chemotyping and MYs production.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

5

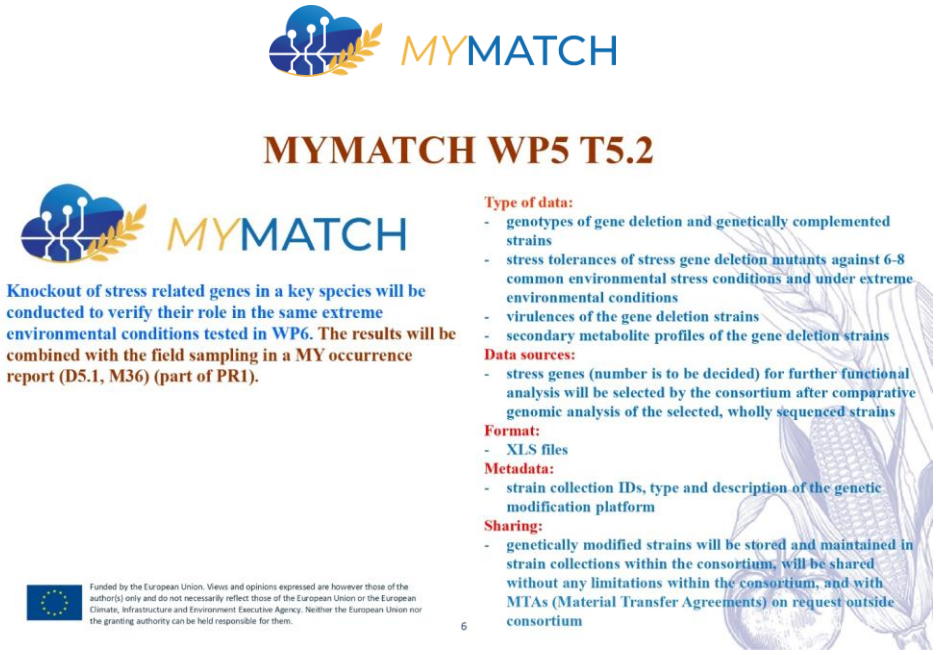
Type of data:
- genome sequences of at least 250 strains (two, at least 1-2 Gbase DNA sequence data files from paired-end sequencing per strain)
- phylogenomic positions of new samples, clonal clusters, major mutations affecting various genes' functions in each genome (*e.g.* secondary metabolite gene clusters), GO term enrichments
- stress protein gene sets determined by orthology search with the *Aspergillus* Stress Protein Database (last version: 19th August, 2024), stress protein family extensions and reductions in comparison to type-strains, GO term enrichments

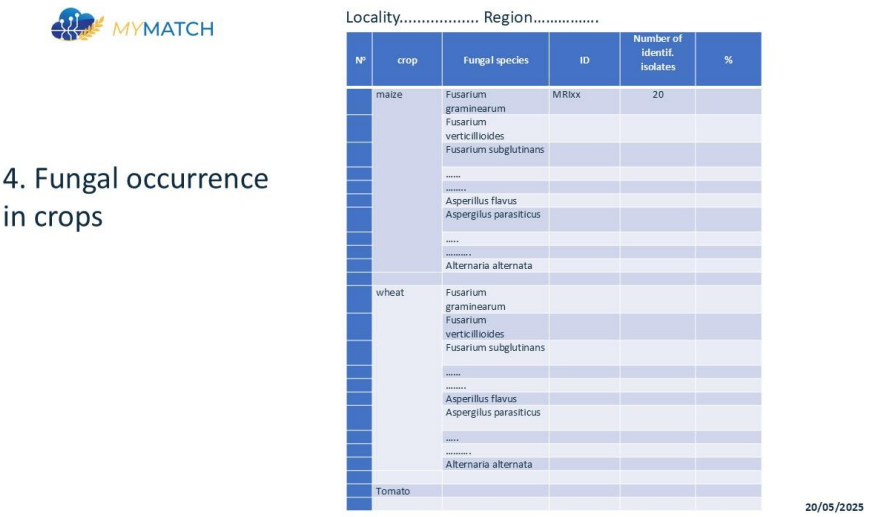
Data sources:
- strains will be selected and sent out for sequencing based on consortial decision

Format:
- FASTQ raw sequencing files, deposited in SRA (Sequence Read Archive, NCBI – NIH, BioSamples with available sequences can be associated with SRA Experiment identifiers and SRA Run identifiers?)
- graphical and XLS files

Metadata:
- SRA metadata, origin of isolate, date of collection, strain collection IDs, sequencing platforms, BioProject and BioSample IDs?

Sharing:
- without limitations within the consortium *via* e-mails, shared cloud space, through SRA, BioProject and BioSample databases?





- In the WP6 session, the presentations provided a full picture of expected data collection activities and data sharing protocols, including clear indications of data granularity and formats.

MYMATCH

MYMATCH

Data Harmonization Seminar

Online Seminar, 9 May 2025

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1

MYMATCH

Sessione 3 - WP6

T6.1: Impact of climate-related factors on fungal ecology and physiology in vitro – Leader: CU

T6.2: Effects of climate-related factors on fungal ecology and physiology in planta – Leader: CU

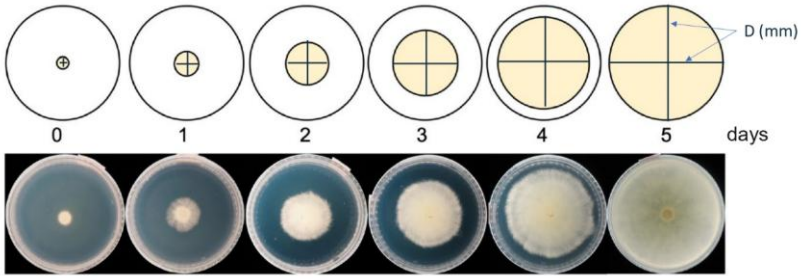
T6.3: Development of functions to update predictive models – Leader: UCSC

MYMATCH

WP6 **T6.1:** Impact of climate-related factors on fungal ecology and physiology **in vitro**

Leader: CU

1- Variable: radial growth rate (μ, mm/day)



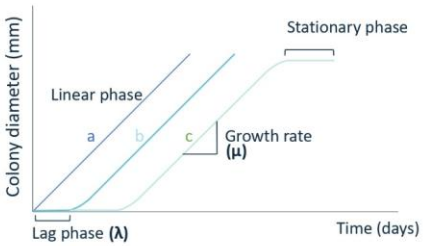
Funded by the European Union

MYMATCH

MYMATCH

WP6 T6.1: Impact of climate-related factors on fungal ecology and physiology **in vitro**
Leader: CU

1- Variable: radial growth rate (μ, mm/day)



Baranyi et al. (1993)

$$y(t) = y_o + \mu_{max} F(t) - \frac{1}{m} ln \left(1 + \frac{e^{m \mu_{max} F(t)} - 1}{e^{m (y_{max} - y_o)}}\right)$$

$$F(t) = t + \left(\frac{1}{\mu_{max}}\right) \ln [\exp(- \mu_{max} t) + \exp(- \mu_{max} \lambda) - \exp(- \mu_{max} t - \mu_{max} \lambda)]$$

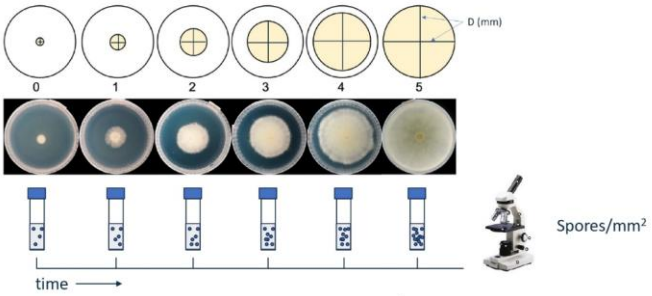
Growth under optimal and suboptimal conditions. **a)** linear model; **b)** linear model with latency phase (lag); **c)** sigmoidal model (lag + stationary pase)

4 09/05/2025

MYMATCH

WP6 T6.1: Impact of climate-related factors on fungal ecology and physiology **in vitro**
Leader: CU

2- Sporulation (Leggieri et al., 2018)



Spores/mm²

5 09/05/2025

MYMATCH

WP6 T6.1: Impact of climate-related factors on fungal ecology and physiology **in vitro**
Leader: CU

1- Variable: radial growth rate (μ, mm/day)

2- Sporulation (spores/mm²)

- Climate Change factors: temperature, water activity, CO_2 levels
- Crops & fungal strains:

MAIZE	WHEAT	TOMATO
Corn Meal Agar	Wheat Agar	V8 Agar
A. flavus	*F. graminearum*	*Alternaria* spp.
F. verticillioides	*Fusarium* spp.	
	Alternaria spp.	

6 09/05/2025

MYMATCH

WP6 T6.1: Impact of climate-related factors on fungal ecology and physiology **in vitro**
Leader: CU

1- Variable: radial growth rate (μ, mm/day)

2- Sporulation (spores/mm²)

3- Mycotoxin production (μg/kg)

- Climate Change factors: temperature, water activity, CO_2 levels

MAIZE		WHEAT		TOMATO	
A. flavus	AFB_1	*F. graminearum*	DON, ZEN	*Alternaria* spp.	AOH
F. verticillioides	FB_1, FB_2	*Fusarium* spp.	DON, ZEN		
		Alternaria spp.	AOH		

7 09/05/2025

Funded by the European Union

MYMATCH

WP6 **T6.1:** Impact of climate-related factors on fungal ecology and physiology **in vitro**
Leader: CU

EXAMPLE

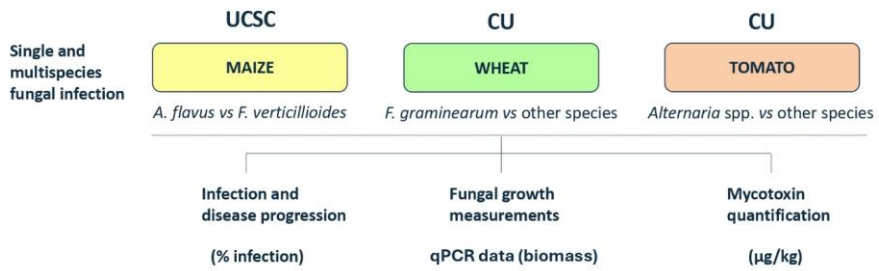
Radial growth rate (mm/day)

Strain	Crop	aw	CO2 (ppm)	Temperature					
				10 °C	15 °C	20 °C	25 °C	30 °C	35 °C
A. flavus	Maize	0.99	400	1.5	2.5	4.1	6.5	8.4	7.2
		0.99	1000	1.3	2.1	3.9	5.8	7.3	6.5
		0.95	400	...	...	...	...	...	...
		0.95	1000	...	...	...	...	...	...

8 09/05/2025

MYMATCH

WP6 **T6.2:** Effects of climate-related factors on fungal ecology and physiology **in planta**
Leader: CU



9 09/05/2025

A major outcome of the seminar was the reaffirmation that a centralized project repository will be developed under WP7. A first version is expected by the end of summer 2025. Each partner will be assigned access credentials for two designated users, allowing secure upload, download, and sharing of structured data. The meeting represented an important step in consolidating a shared understanding of roles, responsibilities, and technical requirements for data handling, in line with the FAIR principles and the project's objectives.

MYMATCH WP5 T5.2

T5.2 – data to be generated and shared

T5.2: Identification and characterization of toxigenic fungal strains isolated from cereals and tomato – Leader: NÉBIH – Participants: MRI, CNR, UNIPR, UMINHO, IFC – M6-M36

(i) Representative strains from the experimental fields will be deposited in culture collections at CNR, NÉBIH, and UMINHO.
(ii) CNR, IFC, MRI, and UMINHO will sequence DNA markers for the taxonomical identification of fungi. Analysis will be set up by CNR, NÉBIH, IFC, and UMINHO.
(iii) They will contribute to the aim separately and partially to reach final objective as consortium, with specific contributes: IFC will be in charge of fungal species identifications and characterizations; NÉBIH will focus on fungal whole genome sequencing, genome assembly, annotation, phylogenomics and stress gene identifications; setting up and maintaining *A. flavus* strain collection will be assigned to IFC; UMINHO will contribute with some genome sequences and fungal species identifications; CNR will contribute with fungal species identifications, comparative genomics, secondary metabolites gene annotation.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1

MYMATCH WP5 T5.2

(iv) Data will converge to identify genetic and metabolic modifications possibly related to adaptation to climate change. This will be obtained comparing selected strains against representative strains isolated over at least a 10-year time frame and available in fungal collection at CNR. In addition, Polyphasic and/or –omics approaches will be used in the mentioned strains (selected from field sampling in MYMATCH and those from fungal collections) to verify fungal response to climate change stress, focusing on genetic traits related to chemotyping and MYs production.

(v) Knockout of stress related genes in a key species will be conducted to verify their role in the same extreme environmental conditions tested in WP6. The results will be combined with the field sampling in a MY occurrence report (D5.1, M36) (part of PR1).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

2

MYMATCH WP5 T5.2

MYMATCH

Agreed

- partner will ship strains at least to 1 culture collection, according to Nagoya protocol.
- DNA markers to sequence:

Aspergillus	Tef, CaM, ITS
Fusarium	Tef, CaM, ITS
Alternaria	GPD, Alta1, ITS

	N° of strains
NFCSO	200
CNR-ISPA	200
UMINHO	Up to 200
MRI	Up to 200
Total	Around 800

CNR, IFC, MRI, and UMINHO will sequence DNA markers for the taxonomical identification of fungi. Analysis will be set up by CNR, NÉBIH, IFC, and UMINHO.

KPI2: B= At the end of MYMATCH, at least 4 genetic traits. A= As multiple genetic traits can be identified by genome sequence analysis, additional traits can be retrieved from genomes, at a rate of 1 every 2 years.

primer name(s) 5'-3'	reference	Aspergillus	Fusarium	Alternaria
ITS1	White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR protocols: a guide to methods and applications. Chapter 38. New York: Academic; 1990. p. 315–22.	X	X	X
ITS4				
Cal1	Glass NL, Donaldson GC. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microbiol. 1995 Apr;61(4):1323–30.	X	X	
Cal2				
Bt2a	Glass NL, Donaldson GC. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microbiol. 1995 Apr;61(4):1323–30.	X		
Bt2b				
Tef1	O'Donnell K, Kistler HC, Cigelnik E, Ploetz RC. Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. Proc Natl Acad Sci U S A. 1998 Jul 21;95(5):2044–9.		X	
Tef1				
Alt1a1	Hong SG, Cramer RA, Lawrence CB, Pryor BM. Alt a 1 allergen homologs from Alternaria and related taxa: analysis of phylogenetic content and secondary structure. Fungal Genet Biol. 2005 Jul;42(7):119–29.			X
Alt1a1				
GPD1	Berbee ML, Pirseyedi M, Hubbard S. Cochliobolus phylogenetics and the origin of known, highly virulent pathogens, inferred from ITS and glyceraldehyde-3-phosphate dehydrogenase gene sequences. Mycologia. 1999 Jan-Feb;91(6):964–77.			X
GPD1				

Type of data:
- isolates identified at sp. level *via* morphological features (around 800) and at sp. level *via* molecular markers (around 400)

Data sources:
- isolates from field sampling, monospore cultures, visual and microscopic observatios, PCR amplicons and DNA sequencing

Format:
- XLS files, containing isolate IDs, species names, DNA sequences (FASTQ), GenBank Accession Nos.

Metadata:
- geographic origin of isolate, date of collection, host, strain collection IDs

Sharing:
- through shared cloud space, GenBank

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

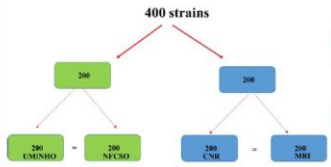
3

MYMATCH WP5 T5.2

MYMATCH

Culture Collections Involved

- CNR (Italy)
- NFCSO (Hungary)
- UMINHO (Portugal)
- MRI (Serbia)



Representative strains from the experimental fields will be deposited in culture collections at CNR, NÉBIH, and UMINHO.

KPI1: B= At the end of MYMATCH, 300 representative strains will be collected totally for the 3 crops. A= As same species can occur on same crop and different countries, multiple representative strains for single species will be collected if detected (rate 20/year).

Type of data:
- selected and characterized representative strains (400) will be deposited in at least two culture collections within the consortium (UMINHO-NFCSO/NÉBIH, CSR-MRI)

Data sources:
- representative strains for single species will be deposited after consortial decision (20/species/year)

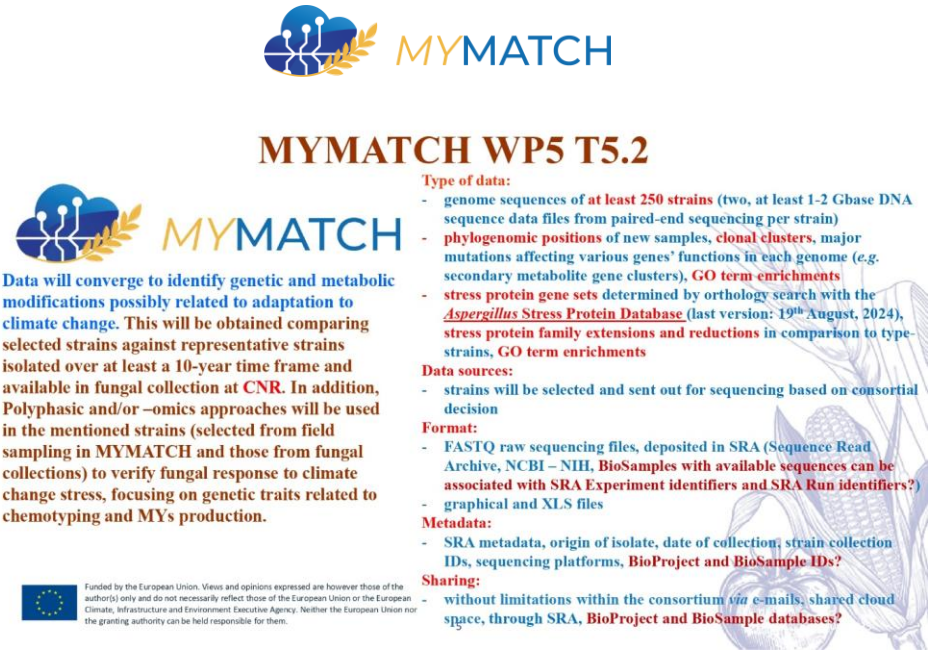
Format:
- XLS files, containing WDCM (World Data Centre of Microorganisms) IDs, authorization to issue strains

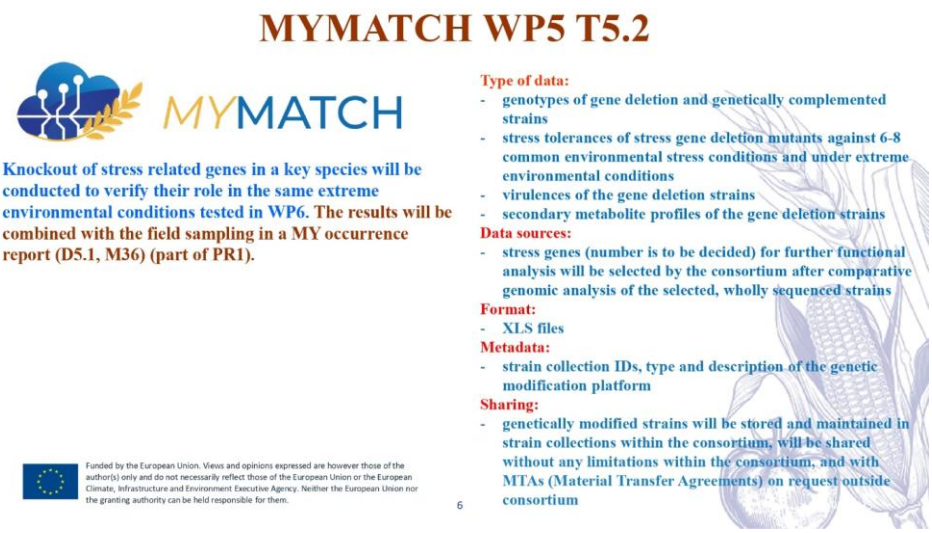
Metadata:
- geographic origin of isolate, date of collection, host, strain collection IDs

Sharing:
- without any limitations within the consortium, and with MTAs (Material Transfer Agreements) on request outside consortium, compliance with and enforcement of Nagoya protocols

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

4





~~8.~~29.2 Data Privacy Notice

Who is collecting your data?

This survey/questionnaire is conducted by Università Cattolica del Sacro Cuore (UCSC), as part of the MYMATCH project funded by the European Union.

For any questions regarding this notice, please contact us at: dpo@unicatt.it.

What data do we collect?

Depending on your responses, we may collect the following personal data:

- No personal information, such as names, emails, phone numbers, etc., will be collected in the proposed surveys
- Climate change perceptions and challenges
- Mycotoxins contamination views and food system awareness
- Digital knowledge
- Opinions and feedback related to MYMATCH (i.e., willingness to engage in multi-actor platform)
- Demographic information (e.g., age, education, gender, location)

Why are we collecting this data?

Your data is collected for the purpose of:

- Informing research or evaluation under the MYMATCH project
- Improving project outcomes or tailoring recommendations
- Reporting anonymized insights to funders or stakeholders

Legal basis for processing

We process your data on the basis of:

- Your consent (by completing the survey and choosing to provide personal data)
- Our legitimate interest in carrying out project-related research, where applicable

How will your data be used?

- Your responses will be anonymized ~~or pseudonymized~~ before analysis
- Personal identifiers (e.g. name or email) will not be shared with third parties
- Only authorized project team members will access identifiable data

How long will we keep your data?

Personal data will be stored securely and retained for no longer than 12 months after the end of the project, unless required by funders or applicable laws.

Your rights

Under GDPR, you have the right to:

- Access your data
- Correct or delete your data
- Withdraw your consent at any time
- Lodge a complaint with your national data protection authority

To exercise your rights, contact dpo@unicatt.it.

Participation is voluntary

Your participation in this survey is entirely voluntary. You may choose not to answer any question or stop at any time.

Thank you

Thank you for your contribution to MYMATCH!

Consent Statement

By answering "Yes" to this form, you confirm that:

- You have read and understood the privacy notice
- You agree to the processing of your personal data as described

~~8.3~~9.3 Declaration of Ethical Commitment and GDPR Compliance

I, the undersigned Paola Battilani in my capacity as Scientific Coordinator of the project *"MYcotoxin*

MAnagement (AI)platform To face CC impact on food safety and Human Health title **MyMatch**",

hereby declare my commitment to comply with the Ethics Rules of Data Processing in accordance

with the **General Data Protection Regulation (GDPR – EU Regulation 2016/679)** and any other

applicable data protection laws relevant to the countries involved in the project.

In particular, I commit to the following principles:

1. **Acknowledgement of ethical responsibility:**

The researchers recognize their responsibility to process personal data ethically and in

compliance with applicable legislation.

2. **Commitment to GDPR and relevant legal frameworks:**

The researchers undertake to comply with the GDPR and with other applicable data protection

laws in the relevant national contexts.

3. **Guarantees of transparency and fairness:**

The project ensures transparency in data processing practices, providing data subjects with

a mis en forme : Justifié

a mis en forme : Justifié

a mis en forme : Justifié

clear, accessible, and understandable information—particularly in the context of surveys and

questionnaires.

4. **Data minimization principle:**

Only data strictly necessary for the achievement of the declared purposes will be collected

and processed, with data retention periods kept to a minimum.

5. **Data security measures:**

Appropriate technical and organizational measures will be implemented to ensure the security

of personal data and to prevent unauthorized access, loss, or improper disclosure (e.g.,

~~pseudonymization,~~ anonymisation, encryption, restricted access, secure backups).

6. **Data subject rights:**

The researchers acknowledge and respect the rights of data subjects, including the right to

access their data, request corrections, and object to processing.

7. **Accountability:**

The researchers accept responsibility for their data processing practices and will take

appropriate steps to address any data breaches or cases of non-compliance.

8. **Ongoing commitment:**

This commitment to ethical data processing is an ongoing process. The researchers will

regularly review and update the project's policies and procedures with the support of the

Ethics Advisor to ensure continued compliance.

a mis en forme : Justifié

a mis en forme : Justifié

a mis en forme : Justifié

a mis en forme : Justifié

a mis en forme : Justifié