AgRimate

D8.2 Data Management Plan 1

July 2025

Funded by the European Union

AgRimate is funded by the European Union under the Grant Agreement No. 101182739. 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 Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

D8.2 Data Management Plan 1

Work package	WP8
Task	T8.4
Due date	31/07/2025
Submission date	31/07/2025
Type of deliverable	Report
Dissemination Level	Public
Deliverable lead	TECNALIA
Version	1.0
Authors	Sara Sillaurren (TECNALIA), Leire Bastida (TECNALIA)
Reviewers	Annekatrin Hoppe, Lina Kasties (UBER); Georgios Ntakos (AUA)
Keywords	FAIR principles, Data security, Sensitive data, Personal data (GDPR compliance), Data anonymization, Informed consent, Data quality assurance

Document Revision History

Version	Date	Description of change	List of contributor(s)
0.1	01/07/2025	Initial ToC	Sara Sillaurren (TECNALIA)
0.2	03/07/2025	First draft	Sara Sillaurren, Leire Bastida (TECNALIA)
0.3	03/07/2025	Updated draft	Sara Sillaurren, Leire Bastida (TECNALIA)
0.4	17/07/2025	Document for review	Sara Sillaurren, Leire Bastida (TECNALIA)
0.5	28/07/2025	Internal review performed	Annekatrin Hoppe, Lina Kasties (UBER); Georgios Ntakos (AUA)
1.0	31/07/2025	Final version	Sara Sillaurren, Leire Bastida (TECNALIA)

Partners

Funding

Funded by
the European Union

Call	HORIZON-CL6-2024-GOVERNANCE-01
Topic	HORIZON-CL6-2024-GOVERNANCE-01-7
Type of action	HORIZON Research and Innovation Actions

Disclaimer

AgRimate is funded by the European Union under the Grant Agreement No. 101182739. 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 Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Copyright Notice

Executive Summary

This deliverable presents the first version of the Data Management Plan (DMP) developed within the scope of the project's commitment to responsible data handling and ethical compliance. The plan outlines the principles and procedures for managing personal and research data generated across the consortium, with particular attention to privacy, security, and regulatory alignment.

The motivation behind this work stems from the increasing complexity of data governance in multi-partner research environments, especially when dealing with sensitive personal data. Ensuring compliance with the European Union's **General Data Protection Regulation** (GDPR) and Horizon Europe's ethics appraisal procedures is not only a legal obligation but also a cornerstone of public trust and scientific integrity.

The methodology followed includes a collaborative assessment of each partner's data management framework, the identification of data flows within demonstration sites, and the definition of protocols for data collection, storage, access, sharing, and deletion. A dedicated **Data Manager** oversees the implementation and monitoring of these protocols to ensure consistency and accountability. In the case of TECNALIA, this role is fulfilled by Tatiana Bartolomé, who is also actively involved in key project meetings—such as the WP Leaders Meetings and data ethics discussions—where she contributes to the coordination and alignment of data management practices across the consortium. As Data Manager, she is responsible for overseeing data governance, ensuring regulatory compliance, maintaining data quality and security, coordinating with project partners, and supporting ethical and **FAIR-compliant data** handling throughout the project lifecycle.

Key findings highlight the diversity of data types and sensitivities across the project, requiring tailored strategies for ethical approval, anonymization, and secure storage. All studies involving human participants are subject to formal ethical clearance from relevant local authorities, in line with the European Commission's codes of practice for research in social sciences and humanities.

The DMP recommends continuous monitoring of ethics issues, regular updates to data handling procedures, and proactive engagement with stakeholders to ensure transparency and adaptability throughout the project lifecycle.

Table of Contents

Abbreviations

CC BY	Creative Commons Attribution
CC0	Public Domain Dedication
CURIE	Compact URI
DMP	Data Management Plan
DOI	Digital Object Identifier
DPO	Data Protection Officer
EDPB	European Data Protection Board
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable
GDPR	General Data Protection Regulation
IP	Internet Protocol
IPR	Intellectual Property Rights
JCA	Joint Controller Agreement
SAL	Security Aspect Letter
T	Task
TCP	Transmission Control Protocol
URI	Uniform Resource Identifier
WP	Work Package

1 Introduction

Within WP8 (Project Management and Coordination), Task 8.4 (Ethics and Data Management) addresses privacy and security concerns related to the use of personal data, in accordance with each partner's operational framework and demonstration environments, and in full compliance with the latest EU GDPR standards. This task includes the development of a comprehensive Data Management Plan (DMP), which defines the scope of the datasets generated and outlines procedures for data storage, access, sharing, and deletion, under the supervision of the appointed Data Manager.

1.1 Aim and Scope

The main objective of this document is to address concerns regarding the privacy and security of various types of data that may be collected, processed, or shared throughout the project. This includes not only personal data—handled in full compliance with the most recent EU GDPR requirements—but also sensitive data, technical and operational data, anonymized or aggregated datasets, and location-based information. All data management practices will be aligned with the specific legal and ethical frameworks of each project partner and tailored to the operational context of their respective demonstration sites.

As part of this effort, a comprehensive Data Management Plan (DMP) will be developed, outlining the nature and scope of the data generated, along with clearly defined strategies for data storage, controlled access, and secure deletion.

All studies involving human participants will strictly adhere to the ethical standards set out in the European Commission's guidance document Ethics in Social Science and Humanities (European Commission, 2021), which provides a robust framework for addressing ethical challenges in SSH research under Horizon Europe. In line with these principles, the project ensures that informed consent is obtained from all participants, with clear and accessible information on study objectives, data usage, and participant rights. Data protection is upheld through GDPR-compliant practices such as data minimization, pseudonymization, and secure storage. Additional safeguards are applied when working with vulnerable populations, including minors or individuals with limited autonomy, and any methodologies involving ethical sensitivities—such as deception, covert observation, or social media data—are subject to rigorous ethical review.

Risk assessments are conducted to identify and mitigate potential harm, including psychological, social, or reputational risks. No research involving human subjects will proceed without prior formal approval from the relevant local ethics committees or institutional review boards, such as the Ethics Committee or equivalent bodies, depending on the demonstration site and national regulations. Oversight of ethical compliance is ensured by the project's Ethics Advisor, which is a representative from TECNALIA, which guarantees that all research activities align with the principles of respect for persons, beneficence, and justice, as emphasized in the European Commission's ethics framework.

This deliverable presents the first version of the AgRimate Data Management Plan (DMP), developed under Task 8.4 to ensure ethical, secure, and FAIR-compliant handling of data throughout the project. In addition to outlining the initial data governance framework, this document sets the foundation for continuous data lifecycle management. A total of four versions of this deliverable will be produced over the course of the project: D8.3 (M21), D8.4 (M42) and a final version (D8.5 M60), at project completion. Each iteration will incorporate updates reflecting the evolution of data collection, processing, and sharing activities, including newly generated datasets, revised data handling procedures, and lessons learned during implementation.

The document begins with an **introduction** that outlines the objectives and scope of the DMP, its relevance to other project tasks, and its importance in ensuring compliance with GDPR and Horizon Europe's ethical standards.

The following sections provide a comprehensive overview of the project's data landscape. **Section 2** offers a detailed summary of the types of data collected, including sensor, video, and human-subject data, along with access, licensing, and preservation strategies. **Section 3** focuses on the implementation of FAIR principles, describing how data will be made findable, accessible, interoperable, and reusable. **Section 4** addresses data security, outlining technical and organisational safeguards, incident response protocols, and long-term protection measures. **Section 5** explores the ethical dimensions of data management, including anonymisation, informed consent, and oversight mechanisms. The document concludes in **Section 6** with a synthesis of key findings and recommendations for future updates. Supporting materials and technical details are provided in the annexes.

1.2 Relationships with other tasks

Task 8.4 plays a central role in ensuring that all data-related activities across the project comply with ethical and legal standards, particularly those involving personal and sensitive information. It supports WP1, WP2, WP3, and WP4 by providing the data governance framework necessary for the responsible development of AR interfaces, robotic systems, and AI-driven analytics. The Data Management Plan developed under T8.4 defines protocols for data collection, storage, access, and deletion, ensuring GDPR compliance across all technical components. This is particularly relevant for tasks involving user interaction and data generation, such as user requirements collection (WP1), the AR Trainer and AR Guide (T2.4-T2.5), robotic pruning systems (T3.1–T3.5), and AI modules like the Pruning Learning Processor (T4.1) and Well-being Analytics Engine (T4.4), which rely on ethically sourced and securely managed data.

In addition, T8.4 is closely aligned with WP5, which focuses on psychosocial and human-centred aspects of the project. The ethical oversight provided by T8.4 ensures that studies assessing psychosocial working conditions (T5.1), peer support systems (T5.2), and human-AI interaction (T5.3–T5.5) are conducted with full respect for participants' rights and well-being. This includes ensuring informed consent, protecting vulnerable groups, and securing ethical approvals from relevant authorities. The collaboration between T8.4 and WP5 is essential to safeguard the dignity, autonomy, and safety of participants, particularly in studies

involving immigrant and female workers, and to ensure that the deployment of AI technologies enhances rather than compromises social sustainability and worker well-being.

Finally, T8.4 is also relevant for WP6, where the demonstration of the solution will be deployed and validated with farmers and workers.

1.3 Importance of the Data Management Plan

A robust and well-structured DMP is fundamental to the success, transparency, and long-term sustainability of the AgRimate project. It serves as a strategic framework for handling all data-related activities, ensuring that data generated, collected, and processed throughout the project is of high quality, securely stored, and readily accessible to authorized stakeholders. By adhering to the principles and procedures outlined in the DMP, the consortium fosters efficient collaboration among partners, facilitates research reproducibility, and ensures full compliance with ethical, legal, and regulatory obligations, including those set forth by the EU General Data Protection Regulation (GDPR) and Horizon Europe's open science policies.

Moreover, the DMP plays a pivotal role in promoting open access to research data, enabling broader dissemination and reuse of results by the scientific community, policymakers, and society at large. It also ensures that data is preserved beyond the project's duration through the use of trusted repositories and standardized metadata practices, thereby contributing to the long-term impact and legacy of the project. Ultimately, the DMP reflects the consortium's strong commitment to responsible data stewardship, scientific excellence, and innovation with societal relevance. It empowers researchers to manage data as a valuable asset, supporting evidence-based decision-making and fostering trust in the outcomes of the AgRimate project.

2 Data summary

A meticulous and comprehensive approach to data management characterizes the AgRimate project. It provides concrete examples from each pilot, precise specifications of formats, data volumes, and collection periodicity, ensuring the integrity and reliability of the project's data management practices.

In all the technical WPs (WP1, WP2, WP3 and WP4), AgRimate project expects to create datasets related to sensors, cameras and data from interviews and surveys, together with data from secondary sources will be the sources for data collecting. As mentioned before, all the ethics requirements when involving human participants will be considered.

In summary, the AgRimate project's data landscape is characterised by:

- **Sensor and Camera data**: High-frequency, high-volume streams from 3D LIDAR, RTK-GPS, IMU, GNSS, audio, and video (RGB and RGB-D images from handheld cameras, 360° cameras, and multi-camera rigs) sensors, primarily in .mp4, wav, .mp3, .ogg, .csv, json, .txt, ROS2, rosbag2 formats. Volumes range from hundreds of MBs to several TBs per day, with semi-annual or activity-based collection. Additionally, there will be information on geospatial data in the way of GPS-tagged vineyard blocks (formats: .gpx, .kml)
- **Human data**: Structured and unstructured data from surveys, interviews, and physiological sensors, stored in .xlsx, .pdf, .csv, .txt and .pptx, mx24 (MAXQDA), and .wma (interview audio) formats; collected at defined intervals tied to pilot activities and research protocols.
- **Annotations**: Visual action labels for video datasets supporting AI and computer vision research, stored in video-compatible formats.
- **Synthetic image data**: composed by Photorealistic synthetic images of vineyards before and after pruning, generated using Blender (not captured by real cameras). The formats are: .jpeg, .png.

Data volumes can be immense, especially for sensor data, with collection periodicity strategically tailored to each pilot's operational and research schedule. This strategic approach instils confidence in the project's planning and execution. This comprehensive and structured approach to data management is not just a process but a crucial element that ensures data from each pilot is managed in a manner that supports the project's objectives, maximises research value, and complies with ethical, legal, and technical standards.

2.1 Access and Licensing

A carefully structured set of procedures governs data access and licensing within the AgRimate project, ensuring that all stakeholders have clearly defined rights and responsibilities regarding data use, sharing, and protection. These procedures are designed to balance the principles of open science with the need to safeguard sensitive and personal information, ensuring that the legitimate interests of all parties involved are respected. AgRimate classifies its datasets into open, restricted, and closed categories, each governed by specific access protocols, licensing terms, and clearly justified limitations where applicable.

Many of the datasets generated by AgRimate—particularly those derived from cameras —will be made openly available to the research community. These datasets will be deposited in trusted repositories like Zenodo, where they are assigned persistent identifiers (DOIs) to ensure long-term accessibility and traceability. Open datasets are typically licensed under permissive models such as Creative Commons Attribution (CC BY) or Public Domain Dedication (CC0), allowing for broad reuse, redistribution, and adaptation, provided appropriate credit is given. Metadata associated with these datasets is also openly licensed (CC0), ensuring that descriptive information remains freely accessible and machine-readable, in alignment with FAIR data principles. For AgRimate project an initial **Zenodo repository** has been created. It can be accessed here: https://zenodo.org/communities/agrimate_eu/settings

However, to be GDPR complaint, datasets containing personal[1] or sensitive information[2]—such as those derived from interviews, surveys, physiological or cognitive measurements, or any data that could potentially identify individuals—are subject to access restrictions. These datasets are accessible only to authorized consortium members or, in specific cases, to external stakeholders with a legitimate interest. This will be managed per case and agreed by all consortia. Access is managed through secure institutional repositories, including private Zenodo collections and controlled environments such as Microsoft Teams or SharePoint, with permissions overseen by designated data managers. These restrictions are grounded in compliance with the General Data Protection Regulation (GDPR) and relevant EU and national data protection laws.

For scientific dissemination purposes, such as publications or conference presentations, data will be processed, aggregated, and/or anonymized to ensure that no individual can be identified. When appropriate, anonymized datasets may be made openly available in accordance with FAIR principles and ethical guidelines, provided that such publication does not compromise data protection obligations or participant confidentiality.

The **AgRimate Joint Controller Agreement (JCA)** formalizes the responsibilities of all partners as joint data controllers, ensuring that personal data is processed lawfully, transparently, and only to the extent necessary for project execution. The JCA also outlines procedures for obtaining explicit consent when processing special categories of data, such as biometric, voice, or image recordings.

Additional restrictions may apply to **datasets protected by intellectual property rights (IPR)**, commercial interests, or contractual obligations. For instance, open dissemination of certain data could compromise a partner's competitive advantage or reveal proprietary technologies. In such cases, access may be temporarily embargoed, with the DMP providing a clear rationale, specifying the legal or commercial basis, the duration of the embargo, and the conditions under which the data may eventually be released. Embargo periods are kept to

[1] https://gdpr-info.eu/issues/personal-data/
[2] https://commission.europa.eu/law/law-topic/data-protection/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en

a minimum and are justified by the need to secure publications, pursue patent applications, or protect innovation potential.

The procedures for accessing both open and restricted data are transparent and robust. Open datasets are made available through standardized, user-friendly protocols and are accompanied by comprehensive documentation and metadata to facilitate discovery and reuse. For restricted datasets, access is granted only after authentication and authorization checks, with audit trails maintained to monitor and verify all access activities. Confidentiality agreements are in place for all personnel handling sensitive data, and all data processing activities are conducted in accordance with the highest ethical and legal standards. In cases where data subjects wish to exercise their rights under GDPR—such as access, rectification, or erasure—a designated contact point within the consortium coordinates the response.

In summary, AgRimate's approach to data access and licensing is grounded in openness and transparency, while maintaining a rigorous commitment to privacy, data protection, and the legitimate interests of all partners. When possible, open data is licensed under CC BY or CC0, with metadata always openly available under CC0. Restrictions are applied only when necessary and are clearly documented in the DMP. This approach ensures that data sharing is maximized for scientific and societal benefit, while fully respecting ethical and legal obligations, reinforcing the consortium's dedication to responsible data stewardship.

2.2 Reuse and Preservation

The AgRimate project adopts a structured and forward-looking approach to data reuse and digital preservation, ensuring that research outputs remain valuable for scientific, societal, and economic purposes. This strategy reflects the consortium's shared commitment to responsible data stewardship and full compliance with legal, ethical, and contractual obligations.

AgRimate's reuse policy is grounded in the FAIR principles, promoting data that is findable, accessible, interoperable, and reusable. The project makes a wide range of datasets—such as robotic sensor streams, annotated video, and processed human factors data—available to third parties, including researchers, industry, and policymakers. Open datasets are deposited in trusted repositories like Zenodo and EOSC, where they are assigned persistent identifiers (DOIs) and enriched with machine-readable metadata. This metadata includes detailed descriptions, provenance, and links to related outputs, and is released under CC0 to ensure broad accessibility.

When necessary, embargo periods are applied—for example, to allow time for scientific publication or to protect intellectual property. These are clearly justified in the Data Management Plan (DMP), kept as short as possible, and followed by open release under permissive licenses such as CC BY or CC0. The embargo periods will be agreed at general assembly level.

Long-term preservation (at least 5 years) is ensured through the use of certified repositories that comply with international standards, such as the OAIS Reference Model and European

Open Research standards[3]. All datasets are accompanied by comprehensive documentation, including data collection methods, processing workflows, and required tools. The information of the **different datasets has been collected** from the partners using the DMP Dataset Survey template in Annex A: DMP Dataset Survey Template. Where project-specific vocabularies are used, they are aligned with common standards to ensure interoperability.

Based on the standardized **dataset survey template** distributed among AgRimate partners, a consolidated summary table has been compiled to capture the **expected datasets generated** across the various Work Packages (WPs) and tasks of the project (a total of 18 datasets for the moment). This summary, presented in Annex B: DMP Dataset Summary Table, provides a structured overview of each dataset, including its origin, responsible partners, timeline, and a brief description of its content and purpose. The information has been extracted directly from the completed partner questionnaires and reflects the current status of data collection and generation activities within the project. As the project progresses, **this information will continue to evolve**. The summary table will be updated accordingly in future versions of the Data Management Plan (DMP) to reflect new datasets, changes in scope, or refinements in dataset descriptions. This effort ensures traceability, facilitates data management planning, and supports alignment with FAIR data principles and Horizon Europe open science requirements.

Quality assurance is integral to the preservation process. Data undergoes thorough review and curation before publication, and the DMP is regularly updated to reflect new data types, evolving standards, and lessons learned. Oversight is provided by the Data Manager and ethics committee, who ensure that preservation practices remain robust, and that data remains accessible and usable over time.

In summary, AgRimate's data strategy ensures that high-quality, well-documented datasets are preserved and shared in a way that maximizes their long-term value, while respecting ethical and legal responsibilities.

2.3 Quality and Monitoring

A robust framework of indicators, systematic review procedures, and clearly defined roles and responsibilities underpins quality assurance and monitoring within the AgRimate project. This structure ensures that all data generated and managed throughout the project lifecycle adheres to the highest standards of integrity, reliability, and compliance.

Key indicators of data quality include completeness, accuracy, consistency, timeliness, and conformity with established standards and norms. These indicators are continuously monitored to detect anomalies or deviations from expected patterns. Both automated validation tools and manual inspections are employed regularly to verify that datasets are comprehensive and free from errors. Cross-checks are conducted to identify inconsistencies

[3] https://open-research-europe.ec.europa.eu/for-authors/data-guidelines/#approvedrepositories

or duplicates across datasets and over time. Additionally, the timeliness of data collection, processing, and availability is assessed to ensure alignment with project timelines and milestones.

The review process is multi-phased and ongoing. Each partner is responsible for internally validating the initial stages of data collection and processing to ensure accuracy. Subsequently, TECNALIA, which oversees data management and quality assurance for the consortium, conducts regular audits and quality assessments. At a minimum, TECNALIA schedules audits and quality checks at three key stages: (i) a preparatory audit before any data collection begins, (ii) in-process audits at agreed intervals while data are being gathered, and (iii) a post-processing audit once each dataset is finalised. The exact timing and recurrence of these checks are adapted to each partner's workflow and the specific risk profile of the dataset. To reinforce this process, a dedicated slot will be reserved in the periodic WP Leaders Meetings to specifically address data quality, privacy, and ethical compliance issues. This will ensure continuous alignment across work packages and facilitate early identification and resolution of potential concerns.

The Data Management Plan (DMP) is a **living document**, reviewed and updated at key project milestones—such as months 21, 42 and 60—and whenever necessary, to incorporate new data types, evolving standards, and insights gained during implementation. Each dataset is accompanied by extensive metadata and documentation, which are also reviewed for completeness and correctness. The ethics advisor, together with TECNALIA team, ensures compliance with ethical standards, the European Code of Conduct for Research Integrity, and all relevant regulatory frameworks. This committee oversees all research involving human participants, ensuring strict adherence to data protection, anonymisation, and informed consent protocols, including [specific regulations or procedures].

Responsibility for quality control and monitoring is distributed among several actors. TECNALIA leads the management and assurance of data quality, including the curation, storage, and preservation of datasets, as well as the enforcement of data protection and anonymisation policies. The **Joint Controller Agreement (JCA)** and the Data Management Plan (DMP) define the technical and organisational measures that each partner must implement when collecting or generating data. Each partner is also accountable for conducting an initial quality check of their respective data. TECNALIA serves as the primary contact for data protection matters, coordinating responses to data subject requests and liaising with supervisory authorities when required. The ethics advisor ensures that all data management activities comply with ethical, legal, and regulatory obligations, thereby providing stakeholders with confidence that their data is handled with the utmost care and respect.

Regular audits and evaluations are conducted to verify compliance with project responsibilities and to identify areas for improvement. These include monthly coordination meetings where data protection, ethical considerations, and data quality issues are systematically reviewed. Particular attention is given to the evaluation of deliverables that may involve sensitive or personal data, ensuring they are aligned with GDPR and ethical standards before submission. The granting authority retains the right to perform external audits, and beneficiaries are

expected to fully cooperate by providing accurate and complete information upon request. Any data breaches or security incidents must be reported within 48 hours, with immediate actions taken to mitigate risks and prevent recurrence.

In conclusion, the AgRimate project adopts a comprehensive and proactive approach to **quality assurance and monitoring**. Through well-defined indicators, continuous review processes, and clearly assigned responsibilities, the project ensures that all data is managed to the highest standards—enhancing its credibility and ensuring the reliability and value of its research outcomes.

2.4 Security and Privacy

The AgRimate project implements a comprehensive set of technical and organisational safeguards, along with a detailed incident response strategy, to ensure the protection of all data—particularly **personal and sensitive information**. These measures are fully aligned with the General Data Protection Regulation (GDPR) and relevant national and EU data protection legislation.

To secure data and systems, the consortium employs both logical and physical access controls. **Logical controls** include user authentication mechanisms, role-based access permissions, and routine audits of access logs to ensure that only authorised personnel can access sensitive information. **Physical security measures**—such as restricted access to facilities, keycard systems, and, where appropriate, biometric verification—are in place to prevent unauthorised entry to data processing environments. Network segmentation is used to isolate sensitive data processing from general network traffic, thereby reducing the risk of unauthorised access or lateral movement in the event of a breach.

All devices and servers are equipped with up-to-date anti-malware solutions that are regularly patched and updated to defend against emerging threats. Secure backup protocols are also in place, including off-site storage and routine recovery drills, to ensure data availability and resilience in the event of accidental loss or cyber incidents. Special category data is subject to enhanced protection measures, such as storage in segregated environments with stricter access controls and comprehensive logging of all access events. Backup copies are stored in separate, secure locations with access limited to a small number of authorised individuals.

The data protection framework is supported by the active involvement of each partner's **Data Protection Officer (DPO) or Privacy Officer**, who is responsible for ensuring compliance with data protection obligations. The **AgRimate Joint Controller Agreement (JCA)** clearly outlines the shared responsibilities of all partners in processing data in accordance with GDPR principles, including lawfulness, transparency, data minimisation, and accuracy. Data processing is strictly limited to the purposes defined within the project scope, and any further use or transfer of data requires a valid legal basis. All personnel handling personal data are bound by confidentiality agreements and receive regular training to reinforce their responsibilities under data protection law.

Anonymisation and pseudonymisation techniques are systematically applied, particularly when dealing with sensitive data. In line with European Data Protection Board (EDPB) guidelines, methods such as data masking, generalisation, perturbation, and pseudonymisation are used to ensure that individuals cannot be re-identified from the datasets. Access to special category data is tightly restricted, with access lists reviewed and updated regularly. Whenever feasible, data is anonymised prior to long-term storage or external sharing, significantly reducing the risk of re-identification.

The incident response protocol is clearly defined and rigorously followed. In the event of a security breach or suspected data incident, the affected partner must notify the consortium without delay and no later than 48 hours after detection. The response process includes immediate containment of the breach, assessment of its scope and impact, and, where required under GDPR, notification to the relevant supervisory authorities. A designated contact point coordinates all communications with authorities and affected individuals. Each incident is thoroughly documented, including the nature of the breach, its consequences, and the corrective actions taken. Post-incident reviews are conducted to identify root causes and implement improvements. Regular audits and simulation exercises are carried out to ensure the continued effectiveness of the incident response plan and the preparedness of all partners.

In summary, AgRimate adopts a proactive and comprehensive approach to data privacy and security. Through advanced technical safeguards, stringent organisational procedures, and a well-established incident response framework, the project ensures the confidentiality, integrity, and availability of all data—while fully upholding the rights and freedoms of data subjects in accordance with GDPR and related regulations.

2.5 Governance and Updates

The governance structure and data management update process in the AgRimate project are grounded in clearly defined responsibilities, a structured review timeline, and robust procedures. These elements ensure ongoing compliance, adaptability to evolving requirements, and alignment with ethical and legal standards. Central to this framework is the appointment of a dedicated **Data Manager**, a qualified professional responsible for overseeing all aspects of data governance. This role includes implementing data policies, monitoring adherence, and ensuring that all consortium members comply with the provisions outlined in the Data Management Plan (DMP). The Data Manager also ensures that necessary adjustments are made in response to changes in project scope or regulatory developments.

Supporting the **Data Manager** (Tatiana Bartolomé) is the **project's Ethics Advisor**, Mr. Jorge Garcia, from TECNALIA. This advisor plays a critical role in upholding ethical standards, particularly in research involving human participants, and ensures that all data-related activities conform to the European Code of Conduct for Research Integrity, the GDPR, and other applicable legal frameworks. Their oversight of sensitive data handling, anonymisation practices, and the protection of data subject rights reinforces the ethical integrity of the project.

The DMP is treated as a **dynamic document**, subject to regular review and revision. In accordance with Horizon Europe guidelines, the initial version is submitted within the first six months of the project. Subsequent updates are scheduled at key milestones and as needed thereafter, particularly when new data types are introduced, processing methods evolve, or regulatory and ethical standards are updated. This iterative review process ensures that the DMP remains current, accurate, and reflective of the project's operational realities, thereby maintaining stakeholder confidence in its relevance and effectiveness.

To facilitate effective coordination, all partners have established permanent communication channels for the timely exchange of information related to data protection, incident response, and compliance matters. For example, Tatiana Bartolomé at TECNALIA serves as the primary contact for **data subject requests** and liaises with supervisory authorities on data protection issues. This approach ensures that all partners remain informed of their responsibilities and any procedural updates, fostering a culture of transparency and shared accountability.

In summary, the AgRimate project's data governance and update process is built on the foundation of a dedicated **Data Manager**, the oversight of an **Ethics Advisor**, a structured DMP review schedule, and continuous communication among partners. Together, these measures ensure that data management remains resilient, responsive, and fully compliant with the highest standards throughout the project's duration.

2.6 Dissemination

The AgRimate project has developed a comprehensive dissemination strategy to ensure that its research findings, data, and overall outcomes reach the appropriate audiences effectively. Recognising the importance of engaging stakeholders through diverse communication channels, the strategy is designed to foster awareness, participation, and impact. Rooted in the principles of open science, transparency, and inclusivity, the plan aims to maximise visibility and accessibility of results across the scientific community, industry, public sector, and broader society.

Dissemination efforts are structured around several key pathways. Open-access platforms such as **Zenodo** and the **European Open Science Cloud (EOSC)** will be evaluated to be used to publish datasets, research outputs, and supporting materials, ensuring global accessibility and citability. Peer-reviewed publications remain a cornerstone of scientific dissemination, and the consortium is committed to publishing in open-access journals whenever feasible, in line with funder requirements and to broaden reach. The project website serves as a central, regularly updated platform offering news, public deliverables, and resources tailored to both expert and general audiences. Additionally, the consortium produces policy briefs, public reports, and white papers to present findings in formats accessible to decision-makers and industry professionals. All **public deliverables** and **outputs of interest to the scientific community** will also be made available through the project's official repository and, when appropriate, published on the Open Research Europe platform to ensure broader visibility and open access. Active participation in scientific

conferences, industry workshops, webinars, and public engagement events further supports direct interaction with a wide range of stakeholders.

A targeted approach is taken to identify and prioritise dissemination audiences. **Open datasets**, **academic publications**, and **technical documentation** are directed toward researchers, including academics, students, and domain experts, to support further research and innovation. Stakeholders—such as technology developers and service providers in agriculture—are engaged through tailored materials, demonstrations, and participation in sector-specific events. Policymakers and regulatory bodies are reached through policy briefs and involvement in relevant forums, supporting evidence-based decision-making and standard-setting. For the general public, the project website features accessible content, and outreach is supported through media engagement and public events to raise awareness of the societal benefits of human-robot collaboration technologies.

To ensure broad outreach, AgRimate employs a mix of traditional and digital communication tools. These include press releases, newsletters, social media campaigns, and printed materials such as brochures and flyers distributed at events. The consortium also leverages partner networks and local dissemination efforts to extend reach across regions and sectors. All dissemination activities are coordinated to maintain message consistency and ensure compliance with ethical and legal obligations, particularly regarding personal data protection as outlined in the Joint Controller Agreement.

The **dissemination strategy** is continuously monitored and refined throughout the project's duration. Key performance indicators—such as the number of publications, dataset downloads, event participation, and media coverage—are tracked to assess effectiveness and inform adjustments. All partners actively contribute to dissemination efforts, with clearly defined roles and responsibilities to ensure broad engagement and long-term impact beyond the project's conclusion.

In summary, the AgRimate dissemination approach is wide-ranging and strategically designed. It integrates open-access publishing, academic and industry engagement, digital and traditional media, and direct stakeholder interaction. This ensures that project outcomes are widely shared, easily accessible, and valuable to all relevant audiences, thereby enhancing the scientific, industrial, and societal impact of the project.

3 FAIR Data

The AgRimate project is firmly dedicated to implementing the FAIR data principles—**Findable, Accessible, Interoperable, and Reusable**—across all its research outputs. This commitment ensures that the data generated and utilised throughout the project is structured in a way that facilitates easy discovery, straightforward access, seamless integration with other datasets, and broad usability for diverse research applications.

3.1 Making Data Findable

The AgRimate project is committed to ensuring that its research outputs are discoverable, accessible, and reusable by adopting persistent and unique identifiers such as **DOIs, URIs, and CURIEs**. These identifiers will be systematically linked to rich, descriptive metadata that provides detailed information about the datasets and research results, thereby enhancing their visibility and facilitating reuse. In alignment with open science principles and the requirements of the funding call, the project will provide access to data necessary for validating scientific publications and will deposit research outputs in open-access repositories under appropriate licenses, particularly in cases of public emergencies.

The project will strictly adhere to a comprehensive set of disciplinary and cross-cutting standards, including:

- **Ethical Standards:** All activities will be conducted in accordance with the highest ethical principles and relevant EU, international, and national legislation. This includes respect for individuals, communities, ecosystems, and cultural heritage, as well as accountability throughout the research lifecycle—from planning and execution to publication and impact.
- **Research Integrity:** All beneficiaries will comply with the European Code of Conduct for Research Integrity, which promotes reliability, honesty, and transparency in conducting, reviewing, and communicating research.
- **Data Protection:** Personal data and open science considerations will be addressed in the Data Management Plan (DMP), which will outline the guiding principles for data handling during and after the project.
- **Intellectual Property Rights (IPR):** The project will follow established rules regarding ownership, access, and use of intellectual property, including both background and foreground results.
- **Gender Equality:** The consortium will actively promote gender balance and equal opportunities throughout the project's implementation, ensuring representation at all levels.
- **Recruitment and Working Conditions:** The project will apply the principles of the European Charter for Researchers and the Code of Conduct for the Recruitment of Researchers, ensuring fair, transparent hiring practices and supportive working environments.
- **Financial and Audit Standards:** Financial reporting will be supported by certificates issued by independent public officials or external auditors, in accordance with standards comparable to those in EU Directive 2006/43/EC.
- **Classified Information:** If EU classified information is involved, it will be managed according to the Security Classification Guide (SCG) and Security Aspect Letter (SAL), following procedures agreed with the granting authority.
- **Legal and Regulatory Compliance:** The project will be executed in full compliance with the Grant Agreement, call conditions, and all applicable legal frameworks at the EU, national, and international levels.

These standards collectively ensure that the project is conducted with **integrity, transparency, and full respect for ethical and legal obligations**, while promoting excellence in research and innovation.

To further enhance discoverability and reuse, the project will embed relevant search keywords within its **metadata**. This metadata will be designed to be open, machine-readable, and aligned with **FAIR principles**, enabling automated systems to harvest and index the information effectively. Metadata will be published under a **Creative Commons Public Domain Dedication (CC0)** or an equivalent license, ensuring unrestricted access and facilitating the broad reuse of research outputs and datasets.

3.2 Making Data Accessible

3.2.1 Repository

The **AgRimate** project will deposit its datasets in **Zenodo**, a reputable open-access repository operated by CERN, recognised for its robust infrastructure and adherence to established data management standards. Arrangements have been made to ensure that Zenodo fully supports the project's needs in terms of secure storage, accessibility, and alignment with the FAIR data principles. Each dataset uploaded to Zenodo will be assigned a unique **Digital Object Identifier (DOI)**, which provides a persistent and citable reference, ensuring long-term access to the associated digital resources.

3.2.2 Data

Not all data generated within the **AgRimate** project will be openly accessible. The project recognises several legitimate scenarios in which data must be restricted or shared under controlled conditions. These limitations fall into two main categories:

Legal and Contractual Constraints

- **Data Protection and Privacy**: Compliance with the General Data Protection Regulation (GDPR) and other relevant data protection laws necessitates careful handling of personal data to prevent breaches of privacy. The consortium will formalise a Joint Controllership Agreement that allocates GDPR responsibilities among partners and includes a consolidated register of each organisation's designated Data-Protection Officer.
- **Intellectual Property Rights (IPR)**: To protect the potential for commercial exploitation and safeguard the legitimate interests of beneficiaries, certain results—including datasets—may be subject to protection for a defined period and territorial scope.
- **Confidentiality and Security**: Sensitive or classified information may require confidentiality, thereby limiting its availability for open dissemination.

Deliberate Restrictions

- **Commercial Sensitivity**: Data may be withheld from open access if its disclosure could compromise the commercial interests of beneficiaries or the competitive position of the EU.
- **Open Access Limitations**: In cases where open access is not feasible, the reasons must be clearly justified in the Data Management Plan (DMP), beyond legal or contractual obligations.

- **Consortium Agreements**: Internal agreements may define specific rights and obligations concerning background and results, which could impose restrictions on data sharing.
- **Ethical Considerations**: Tasks involving ethical concerns require prior approvals, and in some cases, ethical imperatives may necessitate limiting access to certain datasets.

In collaborative projects, individual beneficiaries may choose to restrict access to their data if sharing would conflict with their legitimate interests or with provisions outlined in the Grant Agreement. This approach ensures that legal, ethical, and commercial considerations are respected while maintaining compliance with overarching project obligations.

Where embargoes are applied—for example, to allow time for publication or to secure intellectual property rights—these must be clearly justified in the DMP. The rationale and duration of the embargo should reflect the beneficiaries' legitimate interests, such as protecting results for commercial use or meeting legal requirements. The embargo period should be kept to a minimum to balance protection with the goal of enabling data reuse and transparency.

The DMP will specify the conditions under which embargoes apply, ensuring alignment with open science principles and the expectations of the funding programme. During the embargo, beneficiaries must retain sufficient intellectual property rights to meet future open access obligations. Once the embargo ends, data must be made openly available under licenses such as **Creative Commons Attribution (CC BY)** or equivalent, ensuring broad reuse.

All data will be made accessible through standardised, free-to-use access protocols. The **AgRimate** consortium is committed to providing immediate open access to peer-reviewed scientific publications via trusted platforms such as **Open Research Europe** and the **European Open Science Cloud (EOSC)**. Deposited datasets will be shared under licenses that support wide reuse, in line with the FAIR principles.

Access to data—both during and after the project—will follow the specific terms outlined in the Grant Agreement and its annexes. These provisions aim to balance open access with the protection of intellectual property and beneficiaries' interests.

During the project:

- Beneficiaries will have royalty-free access to results necessary for project implementation.
- Access to background and results for exploitation will be granted under fair and reasonable conditions, unless prior restrictions have been communicated.

After the project:

- Access to results for exploitation purposes will continue under fair and reasonable terms, typically for up to one year after project completion.
- Entities under the same control as a beneficiary may be granted access rights, provided they are identified in the consortium agreement.
- EU institutions, national authorities, and the granting authority may access results royalty-free for policy-making purposes, limited to non-commercial and non-competitive use.

Where restrictions apply—such as for proprietary data or to protect legitimate interests—the DMP will outline the principles governing data access. Even in such cases, the project will strive to uphold the FAIR principles to the greatest extent possible.

For data not suitable for open sharing, the DMP will define the access conditions, ensuring that such data remains available for appropriate use in line with the project's objectives and applicable legal and ethical standards.

To safeguard sensitive data, the project will implement a robust identity verification and access control framework, including:

- **Authentication**: Users must verify their identity using secure credentials (e.g., usernames and passwords).
- **Authorization**: Access rights will be granted based on user roles and responsibilities, ensuring only authorised individuals can access specific datasets.
- **Audit Trails**: Detailed logs will record all access events, including user identity and timestamps, to support monitoring and accountability.
- **Confidentiality Agreements**: Personnel handling personal data will be bound by confidentiality obligations to reinforce data privacy and security.
- **Ethical and Legal Compliance**: All access and data handling will strictly adhere to ethical standards and legal requirements, particularly those set out in the GDPR.

By implementing these safeguards, the **AgRimate** project will ensure secure, compliant, and responsible access to data, protecting both individual privacy and the legitimate interests of all stakeholders.

3.2.3 Metadata

The **AgRimate** project will ensure that all metadata associated with its research outputs is openly accessible and licensed under the **Creative Commons Public Domain Dedication (CC0)** or an equivalent license, as required by the Grant Agreement. This approach is fully aligned with the FAIR principles, particularly supporting machine-actionability, which enables automated systems to retrieve, interpret, and reuse metadata efficiently.

The metadata will include essential elements to facilitate data access and reuse, such as **persistent identifiers (e.g., DOIs)** for each dataset, names of contributing authors, and, where applicable, their institutional affiliations and grant references. This structured and comprehensive metadata framework is designed to enhance the discoverability and usability of the project's research outputs.

In addition, the metadata will provide information on any tools or research outputs necessary to validate the findings presented in scientific publications. This ensures that users can not only access the data but also understand and verify the methodologies and conclusions derived from it.

The project places strong emphasis on the availability of documentation or references for any software required to access, interpret, or reuse the data. This includes detailed information within the repository about relevant tools, models, or instruments developed or used in the project. Where applicable, this will cover software dependencies and usage instructions.

Where possible, software developed within **AgRimate**—including models, algorithms, and designs—will be released as **open-source**, with appropriate licensing (e.g., Apache License) and hosted on accessible platforms such as **GitHub**. This ensures that the software can be freely used, modified, and integrated with the datasets, thereby enhancing their practical value and reproducibility.

However, in cases where software is not released as open-source due to commercial considerations, beneficiaries may **retain intellectual property rights**. In such instances, access to the software—particularly the source code—may be restricted, and any commercial exploitation of project-derived intellectual property may require appropriate compensation to the rights holder.

3.3 Making Data Interoperable

The **AgRimate** project is committed to adopting community-recognised best practices for interoperability, ensuring that both data and metadata can be seamlessly exchanged and reused across disciplines. The consortium will prioritise the use of standardised formats and controlled vocabularies to enhance compatibility and integration with other datasets and systems. Although specific standards are not exhaustively listed, the project's alignment with **FAIR principles** and open science values implies the adoption of widely accepted norms, such as:

- **Data Formats**: Employing commonly used formats like CSV for tabular data, JSON for structured content, and domain-specific formats that support efficient data exchange.
- **Metadata Standards**: Applying established schemas such as Dublin Core, DataCite, or discipline-specific standards that are machine-readable and facilitate discoverability and interoperability.
- **Persistent Identifiers**: Using DOIs, URIs, or CURIEs to ensure consistent referencing and long-term accessibility of datasets and publications.
- **Licensing**: Implementing open licenses (e.g., CC BY, CC0) to support data reuse while respecting intellectual property rights.
- **Trusted Repositories**: Depositing data in platforms like EOSC, which enforce their own interoperability protocols and standards.
- **Open-Source Software**: Publishing models, algorithms, and tools in open repositories under permissive licenses to promote reuse and integration.

In instances where the use of non-standard or project-specific ontologies or vocabularies is necessary, **AgRimate** will provide mappings to widely used ontologies to maintain interoperability and facilitate data reuse. This approach is consistent with the FAIR principles, particularly the emphasis on machine-actionability and semantic clarity.

Any ontologies or vocabularies developed within the project will be openly published, allowing the broader research community to reuse, adapt, or extend them. This practice reflects the project's commitment to open science, including immediate open access to peer-reviewed publications and openly licensed metadata under CC0 or equivalent terms.

The open publication of these semantic resources will enhance the transparency and impact of the project's research, enabling other researchers to better understand and utilise the data produced by **AgRimate**.

Datasets will also include detailed references to related data, whether generated within the project or sourced from previous research or external repositories. The project emphasises the importance of comprehensive documentation within the repository, including any tools, software, or instruments required for data reuse or validation. This structured approach ensures that the data is not only accessible but also understandable and applicable within the broader context of related scientific work.

3.4 Increase of Data Re-use

The **AgRimate** project will implement a range of mechanisms to ensure that documentation supporting data validation and reuse is comprehensive, accessible, and aligned with open science best practices:

- **Metadata**: All metadata associated with deposited datasets will be openly available under a **Creative Commons Public Domain Dedication (CC0)** or an equivalent license. This metadata will include detailed descriptions, deposit dates, author and institutional information, publication venues, embargo periods, funding acknowledgements, licensing terms, and persistent identifiers. It will also reference related publications and outputs to provide essential context for data interpretation and reuse.
- **Repository Deposits**: Datasets will be stored in trusted repositories such as **EOSC**, ensuring open access under licenses that promote broad reuse, such as **CC BY**. These repositories will also host supplementary documentation, including README files, codebooks, and other relevant materials that support data understanding and application.
- **Open Science Commitment**: The project is fully aligned with open science principles, ensuring immediate open access to peer-reviewed publications and making all metadata machine-readable and openly licensed. This commitment extends to the documentation that accompanies the data, reinforcing transparency and reproducibility.
- **Ethical and Legal Compliance**: All data and documentation practices will strictly adhere to ethical guidelines and legal obligations, including full compliance with the **EU General Data Protection Regulation (GDPR)** for the handling of personal data.
- **Quality Assurance**: A rigorous quality assurance framework will be applied to ensure the accuracy, clarity, and completeness of both data and its associated documentation. This includes regular checks to confirm that documentation is user-friendly and fit for purpose.
- **Dissemination and Communication**: The project will actively promote awareness and understanding of its data and documentation through public reports, guidelines, and other communication tools designed to support replicability and encourage reuse.

Through these measures, **AgRimate** will ensure that all documentation necessary for validating data analyses and enabling reuse is openly accessible, well-structured, and compliant with the highest standards of data management and open science.

All datasets will be made freely available in the public domain to support widespread reuse, in accordance with the FAIR principles and the requirements outlined in the Grant Agreement. Metadata will be licensed under **CC0** or an equivalent license, while the data itself will be shared under the most recent version of the **Creative Commons Attribution License (CC BY)** or **Public Domain Dedication (CC0)**, following the principle of "as open as possible, as closed as necessary."

This approach ensures that data is accessible for reuse by the broader community while respecting any legitimate constraints or interests. The project partners are fully committed to open science and will deposit data in reputable repositories to guarantee open access for all interested stakeholders.

The **AgRimate** project is committed to ensuring that the data it generates remains accessible and usable by third parties, even after the project's completion. A series of measures have been established to promote data accessibility and support its exploitation by interested stakeholders:

- **Exploitation of Results**: Beneficiaries are required to make efforts to exploit project results—either directly or through licensing or transfer—for up to four years following the project's conclusion.
- **Horizon Results Platform**: If results remain unexploited one year after the project ends, beneficiaries must publish them on the **Horizon Results Platform** to attract potential users or investors.
- **Open Science Commitment**: The project will follow open science principles, ensuring immediate open access to peer-reviewed publications and making research data available through trusted repositories.
- **Licensing and Metadata**: Metadata will be openly licensed under **Creative Commons Public Domain Dedication (CC0)** or an equivalent license, while datasets will be shared under licenses that allow broad reuse, such as **CC BY**.
- **Standardisation**: If project results contribute to standardisation efforts, beneficiaries must request that the relevant standardisation bodies acknowledge the EU funding.
- **Public Emergency Provisions**: In the event of a public emergency, beneficiaries may be required to grant non-exclusive licenses to legal entities addressing the crisis.

These provisions are designed to facilitate the reuse of data by third parties during and after the project, while also respecting legal, ethical, and commercial constraints such as intellectual property rights and data protection regulations.

The **AgRimate** project also places strong emphasis on documenting the **provenance** of its data using recognised standards. In line with the FAIR principles—ensuring data is Findable, Accessible, Interoperable, and Reusable—the project will ensure that all datasets are accompanied by detailed provenance information to support traceability and reproducibility.

To this end, data will be deposited in trusted repositories such as **EOSC**, under licenses that support open access and reuse. Metadata will be openly licensed (e.g., under **CC0**) and will include comprehensive details such as dataset descriptions, deposit dates, author and institutional affiliations, publication venues, embargo periods, funding sources, licensing terms, and persistent identifiers.

By implementing these practices, **AgRimate** aims to ensure that its research data is not only accessible but also transparent, verifiable, and reliable—thereby enhancing the long-term value and impact of the project's scientific contributions.

The **AgRimate** project has established a robust framework for data quality assurance to ensure the integrity, reliability, and usability of all data generated throughout the project lifecycle. These protocols, outlined across various components of the project documentation, include the following key elements:

- **Ethics Oversight**: The ethics committee, composed of the project coordinator (**TECNALIA**) and the Ethics Advisor from **TECNALIA** will oversee ethical compliance. This body will monitor all ethics-related matters and ensure alignment with applicable ethical standards and guidelines.
- **Adherence to Ethical and Legal Standards**: The project will strictly follow the highest ethical and legal norms, including the **European Code of Conduct for Research Integrity** and other relevant frameworks. Research involving human participants will comply with the European Commission's guidance on ethics in social sciences and humanities.
- **Data Protection**: Personal data will be collected under controlled conditions with informed written consent. All data will be processed lawfully, fairly, and transparently, with appropriate anonymisation techniques applied to recordings and transcripts to protect individual privacy.
- **Data Security**: Both raw and processed data will be stored securely on restricted-access servers to maintain confidentiality and prevent unauthorised access or tampering.
- **Research Data Management**: The project will ensure responsible handling of digital research data in accordance with the **FAIR principles** (Findable, Accessible, Interoperable, Reusable). This includes the development and regular updating of a Data Management Plan (DMP), the use of appropriate repositories, and the promotion of data accessibility and reusability.
- **Open Science Practices**: AgRimate is committed to open science, ensuring that research data and associated metadata are openly available. Metadata will be licensed under **Creative Commons Public Domain Dedication (CC0)** or an equivalent license to support transparency and reuse.
- **Quality Control and Oversight**: **Factor Social** will lead the data management and quality assurance processes, including the curation, storage, and long-term preservation of datasets. **CSIC**, as the dissemination lead, will promote the use of open licenses to facilitate data sharing and reuse.
- **Audits and Compliance Reviews**: The granting authority reserves the right to conduct audits and evaluations to verify compliance with the Grant Agreement. All beneficiaries are required to fully cooperate and provide accurate, complete information during such reviews.
- **Research Integrity**: All project partners must uphold the core principles of research integrity as defined in the European Code of Conduct. These include reliability, honesty, respect for colleagues and society, accountability, and adherence to sound scientific practices.

These comprehensive measures are designed to ensure that the data produced by the **AgRimate** project is of high quality, ethically sound, and suitable for reuse by the broader research community and other stakeholders.

4 Data security

The **AgRimate** project has implemented a comprehensive set of technical and organisational safeguards to ensure the protection of all data throughout the project's duration, including personal and special category data. A key component of this framework is the consortium's operation under a **Joint Controller Agreement (JCA)**, which clearly defines the shared responsibility of all partners in processing data in full compliance with the **General Data Protection Regulation (GDPR)** and any other applicable national or EU legislation.

4.1 Organisational and technical steps

To prevent unauthorised access to data and systems, the **AgRimate** project enforces strict logical access controls. These include user authentication protocols, role-based access permissions, and routine audits of access logs to monitor and verify all access activities. Physical access is equally well protected, with entry to data processing facilities restricted through secure infrastructure, keycard systems, and, where necessary, biometric authentication.

The project also employs robust network security measures such as **network segmentation**, **firewalls**, and **anti-malware solutions**, all of which are regularly updated and patched to address emerging vulnerabilities. Backup and recovery procedures are in place, supported by redundant infrastructure, to always ensure data availability and integrity. Regular system updates and security assessments further strengthen the project's overall cybersecurity posture.

Additional safeguards are applied to **special category data**, such as biometric identifiers, health records, or other highly sensitive information. This data is stored separately from general datasets, protected by enhanced access controls and detailed audit logging. Access to this information is strictly limited to authorised personnel, with access logs maintained automatically and reviewed regularly. Backup copies of such data are stored in multiple secure locations, with physical access restricted to a small number of designated individuals. Furthermore, strict controls are in place to govern the duplication and transfer of this data, including when it is handled via physical media, ensuring that the same high security standards are consistently applied.

4.2 Protocol for Responding to Incidents

The **AgRimate** project has established a clear and prompt incident response protocol to address potential data breaches or security incidents. In the event of a confirmed or suspected breach involving personal data, the affected partner is required to notify the rest of the consortium without delay—and no later than **48 hours** after discovery.

The response process includes immediate containment of the incident, assessment of its scope and severity, and, where applicable under the **General Data Protection Regulation**

(GDPR), notification to the appropriate supervisory authorities. The designated contact for data protection matters—currently **Tatiana Bartolomé** at **TECNALIA**—is responsible for coordinating all communications with regulatory bodies and affected data subjects.

A comprehensive incident log is maintained, documenting the nature of the breach, its impact, and all remedial actions taken. Following the resolution of the incident, a post-incident review is conducted to identify root causes and implement improvements to strengthen the project's overall data security framework.

4.3 Rights of Data Subjects and Compliance

The **AgRimate** project ensures full compliance with the **General Data Protection Regulation (GDPR)** and all applicable national and EU data protection laws. Personal data is processed exclusively when necessary for the execution of project-related tasks, and always in a lawful, fair, and transparent manner. For the processing of **special category data**, explicit and informed consent is required. Data subjects are clearly informed of their rights, including the right to access, rectify, erase, and object to the processing of their personal information.

A designated contact person is responsible for coordinating and managing data subject requests, ensuring that all responses are handled efficiently, appropriately, and in a timely manner.

All personnel with access to personal data are required to sign **confidentiality agreements** and receive regular training on data protection and privacy practices. The **non-disclosure provisions** included in the consortium agreement further ensure that data is only shared or transferred when strictly necessary and supported by a valid legal basis.

4.4 Long-Term Safety and Protection

When a partner discontinues its involvement in the **AgRimate** project or the collaboration comes to an end, data processing activities related to that partner are promptly terminated. The data is either securely deleted or, where necessary, retained in a restricted state for the duration required to address any potential legal liabilities. Once this period has elapsed, the data may be anonymised and preserved for research purposes, provided that appropriate technical and organisational safeguards are maintained in accordance with the **General Data Protection Regulation (GDPR)**.

Data is securely stored and curated over the long term in trusted platforms such as **Zenodo** and **Microsoft Teams**, with access rights strictly managed by designated consortium partners.

Any reuse of AgRimate data for training AI or machine-learning models will (i) rely on a documented GDPR lawful basis and, where relevant, a DPIA; (ii) respect Articles 3–4 of Directive 2019/790 by using only content whose rightsholders have not opted out or for which a licence exists; and (iii) from 2 August 2025 onward, follow the EU AI Act's transparency and data-governance requirements, including keeping a provenance log and, where applicable, publishing the Commission's template summary of training sources.

In summary, **AgRimate** adopts a comprehensive approach to data security, combining advanced technological safeguards, rigorous organisational protocols, and a clearly defined incident response strategy. These measures ensure that all project data—particularly personal and special category data—is handled lawfully, securely, and in alignment with GDPR requirements and recognised best practices in research data management.

5 Ethics and data management

5.1 Key Aspects on Ethics

Ethical and legal considerations are central to shaping data sharing practices within the **AgRimate** project, as outlined in the project's ethics review framework. The following key elements define this approach:

- **Ethical Compliance**: The project is committed to upholding the highest ethical standards and complying with all relevant legal frameworks (Article 14), ensuring respect for fundamental values such as human dignity, freedom, democracy, equality, and human rights.
- **Data Protection**: Full compliance with the **EU General Data Protection Regulation (GDPR)** (Article 15) is a core requirement. This includes lawful, fair, and transparent processing of personal data, with measures in place to ensure data accuracy and relevance.
- **Confidentiality and Security**: In accordance with additional requirements from the granting authority (Article 13), the project enforces strict protocols for handling sensitive information, including obtaining written consent prior to any disclosure.
- **Ethics Committee Oversight**: An ethics Advisor monitors compliance with ethical standards and ensures continuous oversight of ethics-related matters.
- **Ethical Approval**: All research involving human participants adheres to the ethical guidelines set out in the European Commission's *Ethics in Social Science and Humanities* and requires formal approval from relevant local ethics bodies.
- **Anonymisation**: The project ensures the prompt de-identification of personal data in accordance with guidance from the **European Data Protection Board**, maintaining participant anonymity throughout the research process.
- **Right to Withdraw**: Participants retain the right to withdraw from studies at any time without consequence and may request the removal of their data if it has not yet been anonymised.
- **Incidental Findings Policy**: A clear policy is in place to manage unexpected findings during research, with a focus on safeguarding participant well-being.
- **Open Science Principles**: The project is committed to responsible data management in line with the **FAIR principles** and ensures open access to peer-reviewed scientific publications.
- **Dissemination**: The dissemination of results is conducted with due regard for protecting the legitimate interests of all project beneficiaries.
- **Impact Evaluations**: The granting authority may conduct impact assessments, and the project is required to provide relevant information in electronic format upon request.

- **Checks, Reviews, and Audits**: The project is subject to periodic checks, reviews, and audits by the granting authority to ensure ongoing compliance with the Grant Agreement, which may affect the status of the grant.

These measures reflect **AgRimate's** strong commitment to ethical integrity, legal compliance, and transparent, secure data sharing practices, as guided by the ethics review process and the provisions outlined in the Description of the Action (DoA).

5.2 Ethics and Data Management

The responses provided by the **AgRimate** consortium to the **Ethics and Data Management Partner Questionnaire** reflect a well-structured, mature, and ethically sound approach to handling personal and sensitive data. This approach is fully aligned with both European and national legal frameworks, as well as the provisions outlined in the project's **Joint Controller Agreement (JCA)**. It demonstrates the consortium's strong and ongoing commitment to responsible and ethical data management practices. This commitment is further reinforced by the active involvement of the project's **Ethics Advisor**—appointed from TECNALIA—who provides continuous oversight and guidance to ensure that all data-related activities meet the highest ethical standards.

5.2.1 Protocols for anonymization and data protection

The majority of partners within the **AgRimate** consortium have established clear protocols for the anonymisation and pseudonymisation of personal data. Commonly applied techniques include **aggregation**, **generalisation**, **data suppression**, and **pseudonymisation**, all aimed at minimising the risk of re-identification. This risk is continuously assessed and mitigated throughout the data lifecycle. Only anonymised or pseudonymised data will be processed for analytical and dissemination purposes. This is particularly important in the context of **biological and behavioural data** collected during pilot activities, where privacy risks are inherently higher.

5.2.2 Data Protection Officers and Safety Steps

All participating institutions within the **AgRimate** consortium have confirmed the presence of designated **privacy or data protection officers**, who are readily accessible for any data-related inquiries. The consortium has implemented comprehensive **security measures** that encompass both **logical and physical access controls**, **restricted data storage**, **regular audits**, and **robust backup procedures** in the event of a data breach.

These safeguards ensure that data is stored securely—typically on internal servers with limited access—and that contingency plans are in place to maintain data availability and integrity. Collectively, these measures provide stakeholders with confidence that all data is being handled with the highest standards of security and care.

5.2.3 Access, Licensing, and Reuse

Access to data within the AgRimate project is governed by a risk-based approach. Non-personal, non-sensitive, or robot-generated data is typically made openly available through certified repositories such as Zenodo, under permissive licenses like Creative Commons Attribution 4.0 (CC BY 4.0) to encourage broad reuse. In contrast, datasets containing personal or commercially sensitive information are subject to restricted or closed access, with availability limited to authorised individuals under strict conditions.

In certain cases, embargo periods may be applied to protect intellectual property or allow time for publication. These embargoes generally last up to one year, after which the data may be released for wider use. All access restrictions are thoroughly documented, with justifications based on legal, contractual, or voluntary grounds.

Standards and Interoperability

The consortium demonstrates a strong commitment to interoperability by adopting widely accepted data formats and standards. For example, robotic data is stored using the ROSbag format, a recognised standard in robotics research, ensuring compatibility with other platforms and datasets. Metadata follows established schemas such as Dublin Core and DataCite, and persistent identifiers like DOIs and URIs are used to ensure data can be reliably located and cited. Additionally, project-specific vocabularies are mapped to community-recognised ontologies to support data integration and reuse across domains.

5.2.4 Quality Control and Ethics Check

Quality assurance is integrated into every phase of data management. It is standard practice to routinely calibrate sensors, verify data collection methodologies, and conduct periodic reviews. The **AgRimate Ethics Advisor** plays a crucial role in ensuring adherence to ethical principles, research integrity, and legal compliance. This includes reviewing and approving all studies involving human participants, safeguarding participants' rights such as the right to withdraw and to restrict the scope of data collected, and offering guidance on ethical matters that may arise throughout the research process.

Detailed information on **FAIR data principles**, including metadata standards, repository practices, and interoperability measures, is provided in Section 3 of this document.

6 Conclusion

The **AgRimate** project has established a robust, forward-looking, and ethically grounded data management framework that aligns with the highest standards of legal, technical, and scientific excellence. By adhering to the **FAIR principles**—ensuring that data is Findable, Accessible, Interoperable, and Reusable—the consortium maximizes the long-term value, transparency, and reusability of its research outputs.

This first version of the **Data Management Plan (DMP)** reflects a comprehensive and collaborative effort across all partners to define clear protocols for data governance, privacy, security, and ethical compliance. The plan integrates **GDPR-aligned safeguards**, rigorous quality assurance mechanisms, and a structured review process that ensures continuous improvement throughout the project lifecycle.

The DMP also supports the project's open science strategy, enabling the publication of anonymized datasets, metadata, and research outputs in trusted repositories and platforms such as Zenodo and Open Research Europe. This approach not only facilitates scientific dissemination and reuse but also reinforces public trust and accountability.

Importantly, the DMP is conceived as a **living document**. It will be updated at key milestones (M21, M42, and M60) to reflect new data types, evolving standards, and lessons learned from pilot activities and stakeholder engagement. This iterative process ensures that the plan remains relevant, responsive, and aligned with both project needs and external regulatory developments.

By fostering a culture of responsible data stewardship, **AgRimate** sets a benchmark for **ethical and secure data handling in human-robot interaction** research. The consortium's inclusive and transparent approach empowers all stakeholders—researchers, industry, policymakers, and the public—to engage with and benefit from the project's outcomes. Ultimately, this commitment will ensure that **AgRimate's data assets** continue to support innovation, evidence-based decision-making, and societal impact well beyond the project's duration.

Annex A: DMP Dataset Survey Template

AgRimate Data Management Plan: Dataset Survey Questionnaire

Dear Partners, as you already know, the project "**AgRimate** - Transforming Pruning for Smallholder Agriculture with Enhanced Productivity and Working Conditions Using AR and Robotics" must ensure open access to research data and deliver a Data Management Plan which must be updated in the course of the project.

What is the Data Management Plan (DMP)?

The DMP is a document describing the life cycle of all the data sets that will be collected or generated by the project. The DMP must provide the identification elements and the **descriptions of the data sets** and it must include details regarding how the research data will be handled during the project and how they will be preserved after it is completed. It must specify which methodologies and standards will be used in the data creation and management and how and when the data will be shared and made open for re-use. In case parts of the data cannot be openly shared, the DMP must provide motivations. The DMP is a document which evolves during the lifespan of the project. A first DMP must be delivered within the first six months of the project (M6). Updated DMP versions will be delivered during the project duration (M21, M42) and the final DMP version will be released at the end of the project (M60), unless otherwise agreed with the EC project officer. Moreover, new versions of the DMP should be produced whenever important changes in the data or data management policy may occur.

How to fill out the Questionnaire?

The Partners are kindly asked to fill out the attached questionnaire by providing details regarding each data set that will be produced by the project. Each data set has to be described separately. In some cases, it may be necessary to provide an estimation or an intention where the research data have not been collected or generated yet. Once the data are available, the details will be updated.

For each data set you are kindly asked to provide information about:

- data set name, identifier and description - including origin (if collected), scale and use of the data;
- standards and metadata, including methodologies for data collection and typologies and data format;

Guidance is provided for each section and subsection.
We kindly ask you to complete the questionnaire and send it back by **Friday 11th of July 2025**.

For further information and clarifications:

leire.bastida@tecnalia.com (adding in the subject **[AGRIMATE DMP QUESTIONNAIRE]**)

Thank you for your kind attention.

Initial DMP Survey questionnaire

1. Data set reference, name and description

1.A. Data set name

Name: *<Insert here>*

[GUIDANCE: Provide a name or a title for the data set. A data set can be defined as a named collection of data units with the same focus and scope.
For homogeneity reasons, data sets should be named as follows: ***"PROJECT ACRONYM: WPnumber: WP title or short description specifying WP aims: Tasknumber: Task title or short description specifying Task aims: additional information specifying coverage and nature of data (if necessary): version number (optional, in case of revisions to help identifying the updates, especially when depositing in repositories that do not track versioning automatically)"*** *.*
We recommend following similar rules for file naming, adding version number and version date.]

1.B. Data set ID

ID: *<Insert here>*

ID type: *<Insert here>*

[GUIDANCE: Provide an identifier for your data set. Persistent identifiers are generally associated to a data set at the moment of archiving, for dissemination or preservation, in a repository. ***If the data set has no persistent identifier at the moment, just type "not yet available". You could indicate it later.*** *A permanent unique identifier can be used as a reference to a resource even beyond the resource lifetime. It facilitates online tracking and citation.*
In addition, indicate the identifier type ("ID type"), such as for example: DOI, Handle, PURL, URL, URN. ***Check the type of persistent identification used by your repository service****.]*

1.C. Data set origin, WP and current status

1.C.1.Indicate if the data set is: ☐ Collected | ☐ Generated

[GUIDANCE: Collected data sets are data sets that use existing data sources; generated data sets are based on data created by the project. Some data sets may contain both collected and generated data. In this case, select both the options]

1.C.2.Indicate the WP and the task: WP*<Insert here>* | Task*<Insert here>*

1.C.3.Indicate the data set current status (choose one option from the guidance): *<Insert here>*

[GUIDANCE: To indicate the data set current status use one of the following:
1) ***available*** *(i.e. the final version of the data set is completed and ready to be published/has been published)*
2) ***in progress*** *(i.e. at the moment the data are being collected)*
3) **not yet available** *(i.e. the data set is associated to a task/subtask that has not started yet)]*

1.D. Data set creators and contributors

1.D.1.Data set partner creator/s

Partner Creator name: *< partner acronym / name of organization>*

(Fields may be repeated if necessary.)

1

[GUIDANCE: The data sets creators are identified among WP coordinators and research team leaders responsible for the data collection, organization, processing and management.]

1.D.2.Contact Person/s

Contact Person name: *<family, given name>*

Indicated the following details for the Contact Person:

ORCID *(if available)*: *<Insert here>*

Affiliation: *<Insert here>*

e-mail address: *<Insert here>*

(Fields may be repeated if necessary.)

[GUIDANCE: The contact person is a person with knowledge of how to access, troubleshoot, or otherwise address issues related to the data set. Moreover, the contact person must be a person with a stable assignment inside the institution]

1.D.3. Rights Holder/s (*If not yet available, these fields may be filled at later stages.*)

Rights Holder name: *<institution's legal name>*

(Fields may be repeated if necessary.)

[GUIDANCE: The rights holders are usually the public or private institution/s to which belong the data set creator/s OR the institution/s on behalf of which the data set was created. If more than one institution is involved, the rights are shared among them.]

1.D.4. Other Contributor/s (*Involved in the data set production and management.*)

Contributor name: *< partner acronym / name of organization>*

Contact Person name: *<family, given name>*

Contributor type: *<choose an option from the guidance>*

ORCID *(if available and applicable): <Insert here>*

Affiliation *(if applicable)*: *<Insert here>*

e-mail address: *<Insert here>*

(Fields may be repeated if necessary.)

[GUIDANCE:

1) ***Data Collector*** *(such as survey conductors, interviewers…)*
2) ***Producer*** *(person or organization responsible for the form of a media product)*
3) ***Project Member*** *(a researcher indicated in the GA)*

*4) **Researcher** (an assistant to one of the authors who helped with research, data collection, processing and analysis but is not part of team indicated in the GA)*
*5) **Research Group** (the name of a research institution or group that contributed to the data set)*
*6) Other (**specify**)]*

1.E. Data set Description

Provide an abstract of the data set. Include all relevant information listed in the guidance.

<Insert here>

[GUIDANCE: Provide a description of the data that will be generated or collected, the intended goals and to whom it may be useful. Include the following questions:
- *nature of the data (for example: experimental observations; survey results; interview transcripts; simulation data; models; software; diaries; lab notebooks; …);*
- *the data scale (for example the number of the analyzed interviews or samples …);*
- *give information on the existence (or not) of similar data;*
- *if new data are created, explain why no-suitable existing data are available, analyze the gap between previous or current data and yours; describe limits of previous works and how you will improve results in your project;*
- *the data sources (in case it is collected): provide full citation of the data sources;*
- *specify the type of relationship between the data set and its sources (it may be derived, an updated version; it may reuse only a part of the original sources);*
- *describe potential users and how the data can be reused by them.]*

2. Standards and metadata

2.A. Methodologies for data collection or generation, data processing and quality assurance

Provide a description including issues listed in the guidance.

<Insert here>

[GUIDANCE:
- *Describe how the data will be collected or generated and which community data standards (if any) or methodologies will be used.*
- *Describe the unit of analysis and the procedures used to process the data (for example the methods of aggregation of raw data).*
- *Indicate how the data will be managed and organized during the project, mentioning for example back up and security measures, folder structures, file naming conventions, different version control and naming.*
- *Describe procedures for quality assurance that will be carried out on the data at the time of collection or generation, data entry, digitization and revision or validation (e.g. quality control measures can include bias and/or scale of measurement; protocols for recording raw data; controlled vocabularies, code lists and choice lists; computer aided procedures to check data completeness; data cleaning procedures …).]*

2.B. Describe typologies, content and quality of data

Provide a description of data. Indicate whether the data are one or more of the following:

- numerical, textual – graphical, visual or tactile
- created in digital form (born digital) or converted (digitized)
- quantitative o qualitative data
- raw, derived or secondary

- cleaned or processed

<Insert here>

[GUIDANCE:
- *Data may be presented in a tabular, textual, graphical form, or may be audio or video.*
- *Data may be directly created in a digital form or may be later converted, e.g. a digitization of a printable document.*
- *Qualitative data deal with descriptions and cannot be measured; quantitative data is information about quantities that can be measured.*
- *The raw materials are the primary sources of the research process, they represent the records of research or events as first described. Secondary sources are based on primary sources. These sources analyze, describe, and synthesize the primary or original source. 'Derived data' means augmented/enhanced/cooked/adjusted data in analytic data processing.*
- *Cleaned data have undergone a manual process for detecting and removing errors and inconsistencies from data in order to improve their quality. Data processing involves several processes other than simply cleaning data, e.g. validation, sorting, aggregation, analysis, categorizing, etc...]*

2.C. Data format

Indicate formats used for the following, when available and applicable:

Formats used for raw data: *<Insert here>*
Formats used in data processing: *<Insert here>*
Formats used for sharing and preservation: *<Insert here>*

[GUIDANCE: Indicate the file formats used in data collection/ generation and processing. Propriety formats should be converted for reuse and long-term preservation. The most accepted file formats are listed here https://guides.library.uwa.edu.au/c.php?g=325196&p=2178568.]

Thank you for your collaboration!

Annex B: DMP Dataset Summary Table

DATA SET NUM	WP	LEAD	WP PROJECT MONTHS	TASK	TASK leader	Task PROJECT MONTHS	REFERENCE PERSON	CREATOR/s	CONTRIBUTOR/s	DATASET NAME	DESCRIPTION
1	1	TECNALIA	M1-M12	1.1	FBK	M1-M4	Tatiana Bartolome / Ana Belen Ruiz	TECNALIA	UBER, TUB, UPA, AUA	AgRimate WP1 Technical and Operational Roadmap T1.1 – Exploring pruning process and technology perception	This dataset contains qualitative and quantitative data from interviews and questionnaires with stakeholders involved in olive and vineyard pruning in Spain and Greece, capturing their practices, tools, and views on technologies like AR and robotics. It offers unique, user-centred insights into both traditional and tech-assisted pruning, integrating psychosocial, cognitive, and operational dimensions.
2	1	TECNALIA	M1-M12	1.1	FBK	M1-M4	Tatiana Bartolome / Ana Belen Ruiz	TECNALIA	-	AgRimate WP1 Technical and Operational Roadmap T1.1 – Analyzing data from surveys on pruning and technology perception	This dataset builds on previously collected data about pruning practices and perceptions, focusing on structured analysis of survey responses. It includes Likert-scale evaluations of AR and robotic technologies across dimensions such as adaptability, ease of use, reliability, safety, trust, utility, as well as participant effort and demographic profiles.
3	1	TECNALIA	M1-M12	1.1	FBK	M1-M4	Ntakos George	AUA	-	AgRimate WP1 Technical and Operational Roadmap T1.1 - Greek Vineyard Pilot Baseline Data	This dataset contains initial agro-environmental data from the AUA experimental vineyard in Spata, Athens, including GPS-tagged locations, field history, and agro-management practices. It supports pilot use case definition, demonstration planning, and baseline KPI tracking before technology integration.
4	2	XYMBOT	M3-M48	2.2	XYMBOT	M6-M24	Ntakos George	AUA	-	AgRimate: WP2: Tree Digital Passport: T2.2: Worker-Vine Interaction History and Tree Health Performance Records: v1.0	This dataset will be part of the digital passport system, capturing vine-specific data such as pruning history, anonymous worker ID, health indicators, stress factors, and yield across seasons. It will be used to train AI decision tools and enable traceability at the vine level.
5	2	XYMBOT	M3-M48	2.2	XYMBOT	M6-M24	Gerardo Beruvides	XYMBOT	-	AgRimate WP2 AgRimate AR Solutions T2.2 - Digital Product Passport	This dataset supports the development of the Digital Product Passport (DPP) by organizing structured, interoperable data on tree pruning operations, including worker metrics, tree health, environmental conditions, and productivity. It enables traceability, enhances AI and XR module integration, and supports accurate digital reconstructions for training and simulation purposes.
6	3	ROB	M3-M58	3.1	ROB	M3-M48	María Beneyto	ROB		AgRimate WP3 AgRimate Robotics Solutions T3.1 - Navigation dataset	This dataset will support the development of an autonomous pruning robot by capturing multi-sensor navigation data using a custom mobile platform equipped with RGB-D cameras, LiDAR, RTK-GPS, IMU, and wheel encoders. Recorded in ROSbag2 format, it includes 1–5 hours of vineyard field data for testing obstacle detection, row-following, and localization algorithms, offering a unique combination of sensors not found in existing public datasets.
7	3	ROB	M3-M58	3.4	IUVO	M6-M48	Luca Morelli	IUVO	-	WP3 AgRimate Robotics Solutions T3.4 - Exoskeleton mechatronics and wearing data	This dataset will include exoskeleton mechatronics data such as joint position, torque, velocity, and actuator parameters, enabling inference of user biomechanics. It will also contain fitting and alignment settings to ensure proper exoskeleton usage, supporting analysis of both device performance and user interaction.
8	3	ROB	M3-M58	3.4	IUVO	M6-M48	Sara Sillaurren	TECNALIA	IUVO	Workers psychopisiological signals and images	Captures of psychophisiological and video signals from the workers to infer thier cognitive and physical state

DATA SET NUM	WP	LEAD	WP PROJECT MONTHS	TASK	TASK leader	Task PROJECT MONTHS	REFERENCE PERSON	CREATOR/s	CONTRIBUTOR/s	DATASET NAME	DESCRIPTION
9	3	ROB	M3-M58	3.6	FBK	M6-M42	Paul Chippendale	FBK	-	WP3:RoboticSolutions:T3.6:EnvironmentDigitisation:VineyardDataset:v0	This dataset consists of georeferenced images of approximately 1,000 vineyard plants, captured before and after pruning over three years, enabling temporal analysis of individual plant growth. Its main goal is to train a computer vision model for pruning prediction, offering a novel resource for researchers and agritech companies due to its unique longitudinal structure.
10	3	ROB	M3-M58	3.6	FBK	M6-M42	Ntakos George	AUA	FBK	AgRimate: WP3 T3.6 - Drone and Robotic RGB-D Imagery for Vine Health Monitoring and Environment Digitisation v1.0	This dataset contains RGB and RGB-D images captured by drones and handheld or robotic sensors before and after pruning over three seasons. It will be used to reconstruct 3D environments, detect pruning cuts, assess vine vigour, and support decision-making and productivity analysis.
11	4	TAU	M6-M56	4.1	TAU	M13-M56		TAU		AgRimate WP4 AgRimate AI Capabilities T4.1 - Synthetic vineyards pruning dataset	This dataset contains over 5,000 synthetically generated photorealistic images of vineyards before and after pruning, created in Blender following the double Royat style with randomized visual parameters. It is designed to test and pre-train pruning learning processor models (T4.1) through scalable, high-variability image data before real-world datasets are available.
12	5	UBER	M6-M56	5.1	UBER	M1-M12	Annekatrin Hoppe	UBER	TECNALIA	AgRimate WP5 Psychosocial and human-centred approaches T5.1 – Psychosocial working conditions among Farmers in Spain and Greece	This dataset contains qualitative data from semi-structured interviews with olive and vineyard farm owners in Spain and Greece, focusing on psychosocial working conditions during pruning and harvesting. It explores job demands, resources, and the role of technological aids, and includes anonymized transcripts and thematic analyses based on Brown and Clarke's (2006) framework.
13	5	UBER	M6-M56	5.1	UBER	M1-M12	Ntakos George	AUA	-	AgRimate: WP5/WP6: Worker Feedback on AR/Robotic Pruning and Social Sustainability KPIs: T5.1 & T6.4: v1.0	This dataset captures psychosocial and ergonomic assessments from workers using AgRimate AR/robotic solutions. Includes interviews, smartphone micro-surveys, and structured forms covering workload, safety, inclusion, perceived usefulness and well-being. Data will be gender/migrant dis-aggregated and used to evaluate KPIs linked to social sustainability, acceptance and training effectiveness
14	7	F6S	M1-M60	7.1	F6S	M1-M60	Nika Levikov	F6S	-	Agrimate: WP7 C&D, Exploitation Roadmaps & Thematic Open Calls T7.1: Stakeholder Database	This stakeholder database contains structured textual data on companies and individuals engaged in the AgRimate project, including organizational details, contact information (with consent), and engagement records. It is stored in spreadsheet format and relies on internal guidelines for consistency, without external data standards or explicit quality control fields.
15	7	F6S	M1-M60	7.2	F6S	M1-M60	David Erice	UPA	-	AgRrimate WP7 C&D, Exploitation Roadmaps & Thematic Open Calls T7.2 - Workshops Registration	This dataset will consist of individual registration forms completed by event attendees, with the final database updated based on actual participation. It may also include additional data related to attendee engagement and project-related evaluations, though the form and exact content are not yet finalized.
16	7	F6S	M1-M61	7.5	F6S	M30-M60	Nika Levikov	F6S	-	AgRimate: WP7 C&D, Exploitation Roadmaps & Thematic Open Calls T7.5: Open Call 1	This dataset documents open call 1 applicants (startups/SMEs) and their proposals, including organizational, contact, proposal, and evaluation details in a structured tabular format. It supports application tracking and evaluation, with data quality depending on applicant input and internal review processes, but without adherence to external standards.

References

European Commission (2021). *Ethics in Social Science and Humanities*. Available online