Document title: Data Management Plan

Version	Status	Date
1.0	Final	2025-15-05

Author
Pär Erik Martinsson

Contact
par-erik.martinsson@ltu.se, 0920-493637

Deliverable D12.5
Data Management Plan (DMP) V2

Work package leader: Pär Erik Martinsson
par-erik.martinsson@ltu.se

Abstract
This document is Deliverable D12.5 of the Arrowhead fPVN project and presents the updated Data Management Plan (DMP), building upon the initial version (D12.1). It reflects progress made, lessons learned, and refinements to data practices throughout the project lifecycle. The DMP emphasizes the implementation of FAIR principles and addresses ethical, legal, and privacy considerations. Updates include revised data classifications, improved handling procedures, and enhanced metadata standards. As a living document, it supports continuous alignment with Horizon Europe guidelines and evolving project needs.

Chips.JU project 101111977 - Arrowhead FPVN
Project Coordinator: Professor Jerker Delsing | Luleå University of Technology
EUROPEAN PARTNERSHIP

Grant agreement no.	**101111977**
Project acronym	**Arrowhead fPVN**
Project full title	**Arrowhead flexible Production Value Network**
Dissemination level	PU
Due Date	31-05-2025
Date of Delivery	31-05-2025
Deliverable Number	D12.5
Deliverable Name	Data Management Plan v2
AL / Task related	LTU
Author/s	Pär Erik Martinsson,
Contributors	Lama Alkhaked
Reviewer	Jerker Delsing
Keywords	Data management plan, Data sharing, data privacy, fair principles
Abstract	This deliverable presents the updated DMP for the Arrowhead fPVN project, reflecting progress, FAIR alignment, and enhanced data management practices.

Table of contents

1. Introduction

The Arrowhead flexible Production Value Network (fPVN) project aims to establish autonomous and evolvable interoperability among stakeholders in the flexible production value network. It is built on three key pillars: the microservices paradigm, the utilization of major industrially accepted data models, and automatic translation between these data models. The project is expected to have a significant impact on manufacturing productivity and flexibility. This deliverable focuses on the management of the data in the project.

As an effort to enhance the findability, accessibility, interoperability, and reusability (FAIR4) of research data, this deliverable encompasses details related to

- The management of research data during and after the project concludes.
- The nature of data collection, processing, and/or generation.
- The methodologies and standards to be employed.
- Considerations regarding data sharing and open access.
- The approach to data curation and preservation, extending beyond the project's completion.

1.1 Objective of the deliverable

The objective of this deliverable is to provide an updated version of the Data Management Plan (DMP) for the Arrowhead fPVN project, reflecting the progress made and lessons learned since the submission of the initial version (D12.1). This updated DMP outlines the current practices, changes, and refinements in data collection, processing, storage, sharing, and preservation across the project’s lifecycle. The purpose is to ensure continued alignment with Horizon Europe guidelines, address evolving project needs, and reinforce transparency, reusability, and secure management of research and technical data.

In this version, particular emphasis is placed on reinforcing the integration of FAIR principles (Findable, Accessible, Interoperable, and Reusable) throughout all data management processes, and embedding ethical considerations related to data handling, privacy, and research integrity. The deliverable also reflects an iterative review of the original DMP contents, ensuring that updates are grounded in the operational realities. This continuous adaptation helps maintain the relevance, compliance, and usability of the DMP as a living document.

1.2 Report Structure

This deliverable follows a structured and modular format aligned with Horizon Europe guidelines for Data Management Plans (DMPs), while integrating project-specific developments and refinements made since the initial version (D12.1). It begins by outlining the scope and objectives of the updated DMP, followed by a detailed overview of the types of data collected, generated, and processed within the AFPVN project. The structure has been adapted to reflect practical insights gained during project implementation, including reclassification of data types and refinement of data handling procedures.

Subsequent sections provide a comprehensive account of data security and privacy measures, data sharing and collaboration protocols, and the long-term preservation strategy adopted by the consortium. The document also addresses metadata standards, documentation practices, and clearly defined roles and responsibilities to ensure compliance and accountability.

The inclusion of updated FAIR data management strategies underscores the project's commitment to making data Findable, Accessible, Interoperable, and Reusable. In addition, ethical considerations related to data handling, research integrity, and participant privacy have been explicitly incorporated.

The report concludes with mechanisms for continuous review, support, and quality assurance, treating the DMP as a living document that evolves in response to project realities, stakeholder feedback, and regulatory developments. Where applicable, revisions and enhancements from the original D12.1 version are clearly indicated to maintain transparency and traceability of updates.

2. AfPVN Data Summary

The primary purpose of data collection within the Arrowhead fPVN project is to enable the development, integration, and validation of autonomous and evolvable interoperability technologies across diverse production value networks. Data is collected to support core activities such as requirements engineering, architecture design, semantic model alignment, and use case implementation. Specifically, technical data from industrial partners facilitates the testing of the microservice-based framework and automated translation tools, while administrative and management data ensures efficient coordination and compliance monitoring. Moreover, collected data underpins iterative development, performance evaluation, and demonstration of interoperability solutions in real-world scenarios. Ensuring that all data serves a well-defined technical or organizational objective is essential to maintaining compliance with ethical standards, data minimization principles, and the overall goals of the project.

The types of collected/generated data in this project can be roughly categorized into four categories:

- Experimental Data
- Source Code and Technical Artifacts
- Standard and Model Mapping Data
- Administrative and Communication Records
- Metadata and Documentation

2.1 Types of data

The Arrowhead fPVN project involves a wide range of activities spanning the development, integration, and validation of interoperable digital technologies across multiple industrial domains. As a result, diverse types of data are generated, including technical, experimental, semantic, and administrative data. This section provides an overview of the key data types collected and produced throughout the project. Understanding these data types is essential for establishing effective management practices, ensuring compliance with ethical and legal requirements, and facilitating reuse and long-term preservation. Furthermore, this classification supports clear assignment of responsibilities and enables the implementation of FAIR principles throughout the data lifecycle.

Table 1:Data Types

Category	Description
Experimental Data	Resulting from integration, validation, and demonstration activities across domains such as automotive, aerospace, green energy, and

	process industries. These are collected during testing of microservices and semantic translation components.
Source Code and Technical Artifacts	Developed under the microservices paradigm and managed through version-controlled repositories (GitHub)
Administrative and Communication Records	Including meeting minutes, internal reports, and coordination documentation, primarily stored in secured cloud platforms (eOwnCloud)
Metadata and Documentation	Descriptive information about datasets, code, and project results, ensuring data discoverability and adherence to FAIR principles

After identifying and categorizing the various data types generated within the Arrowhead fPVN project, it is essential to outline how these data are securely handled throughout the project lifecycle. The following table provides an overview of the storage solutions used for each data type, the access control mechanisms in place to ensure data confidentiality and integrity, and the backup procedures adopted to prevent data loss. This information ensures transparency and demonstrates the project's commitment to robust data management practices in line with Horizon Europe and FAIR data requirements.

Table 2:Data storage, access control, and backup overview.

Category	Data Storage	Access control	Data backup
Experimental Data	The partner responsible for a particular data will decide the most appropriate storage method, like OwnCloud, GitHub, GitLab	Access and sharing policies for data will be determined on a case-by-case basis by the responsible partner	Data backup and preservation for data will be the responsibility of the respective partner.
Source Code and Technical Artifacts	Code repositories will be maintained on both GitHub and GitLab to ensure redundancy and accessibility	Repositories will be accessible to project collaborators and contributors, with version control and issue tracking	GitHub and GitLab provide automatic versioning, backup, and redundancy
Administrative and Communication Records	Data of this category will be stored on OwnCloud, a secure	Access will be restricted to	Regular backups of administration data

	cloud storage platform.	authorized project personnel	will be maintained on OwnCloud.
Metadata and Documentation	Data of this category will be stored on OwnCloud, a secure cloud storage platform.	Access will be restricted to authorized project personnel	Regular backups of administration data will be maintained on OwnCloud.

2.2 Data Security and Privacy

The Arrowhead fPVN project is committed to maintaining high standards of data security and privacy, in accordance with the General Data Protection Regulation (GDPR) and best practices for responsible data management. While the project does not involve the collection or processing of personal data, all other project data will be handled with strict confidentiality. Access to data will be restricted to authorized personnel, with access permissions carefully managed to ensure that only individuals with a legitimate project-related need can retrieve or modify data. Data will be securely stored on trusted platforms, with encryption applied during transmission and, where appropriate, at rest. Regular backups will be performed to prevent data loss, and secure data transfer methods will be used when sharing information among consortium partners. Though personal data is not processed, the project has still adopted proactive safeguards, including internal review procedures and secure coding practices, to ensure overall system integrity and reduce the risk of data-related vulnerabilities.

2.3 Data Sharing and Collaboration

The Arrowhead fPVN project promotes efficient and transparent collaboration through the use of multiple data-sharing platforms, including GitHub, GitLab, OwnCloud, and other relevant tools as required. A structured collaborative workflow guides how data is created, reviewed, approved, and published among project partners, ensuring consistency and accountability. Version control is implemented using Git to manage code and documentation, enabling contributors to track changes through branches and pull requests. To foster a productive and respectful working environment, a code of conduct has been established and communicated across the consortium, promoting open collaboration and addressing potential conflicts constructively. All shared data will carry appropriate licensing terms; code repositories will specify open-source licenses, and any shared documentation or datasets will include clear licensing information. Before any data is integrated into the project's outputs, it undergoes a review and approval process to ensure quality and relevance. Access to shared data is governed by well-defined policies that outline who can access which data, under what conditions, and for what purposes. Ownership and intellectual property rights are addressed through formal data-sharing agreements, recognizing that some outputs (e.g., code) may be open-source, while others (e.g., experimental data) remain under the stewardship of the responsible partner. Public dissemination of project results, particularly code and documentation, will be facilitated via appropriate online repositories and linked through the project website, with Zenodo suggested

for long-term open access. Provisions are also in place for the retraction or removal of data found to be outdated, inaccurate, or in violation of agreements, and all such actions will be documented. To encourage broader participation, contribution guidelines for external collaborators are provided, clarifying how they may submit code or documentation. An archive of data versions will be maintained to preserve transparency and reproducibility, and research-related data will be stored for a minimum of 10 years. If needed, a Data Access Committee may be established to oversee access requests, particularly for sensitive data. Finally, continuous communication and support mechanisms will ensure smooth coordination across the consortium and timely resolution of any data-sharing issues or disputes.

2.4 Data Preservation

The fPVN project is committed to ensuring the long-term preservation of its data to maintain availability and usability beyond the project's lifespan. Data preservation efforts will follow established methodologies and best practices, tailored to the characteristics of each data type. A dedicated data retention policy will define how long different categories of data will be stored, with retention periods based on data type and relevant legal requirements, including data protection regulations. Archived materials will remain accessible to project personnel and, when appropriate, to the wider community. Access will be governed by defined data access policies, as well as legal and ethical constraints. In Sweden, research data is regarded as public documentation, which can be requested by the public; however, a secrecy assessment is conducted in each case, in accordance with the Public Access to Information and Secrecy Act (https://www.government.se/information-material/2009/09/public-access-to-information-and-secrecy-act/). To ensure long-term accessibility, data will be stored in open and widely accepted formats, and all associated metadata and documentation will be preserved to support future understanding and use. The project will also periodically review data formats and, when needed, migrate or upgrade them to align with technological advancements. Comprehensive backup and redundancy protocols will be implemented across all data types, safeguarding against loss due to hardware failure, accidents, or other unforeseen events. Regular audits will verify the integrity and accessibility of preserved data, enabling early identification and resolution of issues. Procedures will be established for the transition of data preservation responsibilities in the event of staff changes or project conclusion, ensuring continuity and accountability. Finally, data that is no longer required will be securely and compliantly disposed of, in line with legal and regulatory standards.

2.5 Metadata and Documentation

The fPVN project will employ recognized and standardized metadata schemas, such as Dublin Core, to describe and provide context for various types of data. These standards will be adapted as needed for specific data types. Metadata schemas will be chosen to support interoperability and discoverability across different data platforms and repositories. Metadata of deposited data must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles ((Findable, Accessible, Interoperable and Re-usable data)) and provide descriptive information as explained in the sections for the different types of the provided data.

2.5.1 Metadata elements

Metadata will include essential elements that describe each dataset comprehensively, including but not limited to:

- Title and description: Clear and concise titles and detailed descriptions of datasets to convey their purpose, contents, and relevance.
- Creator and contributor information: Names, affiliations, and contact details of those who created or contributed to the dataset.
- Date of creation and modification: The dates when the dataset was created, updated, or modified.
- Data format and structure: Details about the data format, structure, and file types.
- Data source: Information about the source of data, such as instruments, sensors, or data collection methods.
- Licensing and access information: Information about data access rights and usage restrictions, including licensing terms.
- Keywords and subject categories: Keywords and subject classifications that enable effective searching and categorization.
- Version and history: Documentation of dataset version history, updates, and any related datasets.
- Data quality and provenance: Information on data quality, accuracy, and the origin of the dataset.
- Related publications and references: Links to relevant publications, research papers, or documentation that provide additional context.

2.5.2 Documentation

The project will implement clear guidelines for the creation, maintenance, and regular updating of documentation related to different types of generated data. These documentation standards will be adapted to the specific characteristics of each data type, with a strong emphasis on completeness, consistency, and clarity. Code repositories will be accompanied by comprehensive README files containing key information for users and collaborators. This will include installation and setup instructions, descriptions of code structure and organization, usage examples and sample code, lists of dependencies and prerequisites, troubleshooting tips, FAQs, and contribution guidelines for external participants. Administrative data will be documented in a structured and organized manner to support effective project coordination and data management. This documentation will include records of communications and meetings, project reports, progress updates, meeting minutes, and agendas. Additionally, partners responsible for experimental data will ensure detailed documentation relevant to their specific datasets, including data collection methods, preparation processes, and any unique requirements or considerations for the data’s intended use. Through this approach, the project aims to ensure transparency, reproducibility, and ease of collaboration.

2.6 Roles and Responsibilities

Clear and well-defined roles and responsibilities are essential for effective implementation and compliance with the Data Management Plan (DMP) throughout the project's lifecycle.

- Project Coordinator: Oversees the overall DMP implementation.
- Code Maintainers: Responsible for code review, merge, and maintenance.

- Responsible Partners: Ensure proper handling and storage of data.
- Data Manager: Ensures the overall compliance with the DMP.

2.7 Review and Update

The fPVN project will implement a structured schedule for periodic reviews of the Data Management Plan (DMP) to ensure that data management practices remain aligned with project goals and respond effectively to evolving requirements. An initial review timeline will be established, with subsequent reviews conducted annually at regular intervals. All project partners are encouraged to contribute input and feedback during these reviews, promoting inclusive decision-making and incorporating diverse perspectives. The review process will evaluate the effectiveness and compliance of data management practices against several key criteria: adherence to legal and regulatory requirements, consistency with the original goals of the DMP, responsiveness to changing project priorities, data quality and accuracy, accessibility and discoverability of data, the robustness of data preservation measures, and the completeness of documentation for different types of data. Findings from each review will be documented in a formal report, which will highlight strengths, identify gaps, and propose areas for improvement. Feedback will be gathered through mechanisms such as surveys and direct communication channels to ensure ongoing engagement from partners and stakeholders. Based on these findings, a detailed action plan will be developed, outlining the specific measures, responsibilities, and timelines required to implement updates. All revisions to the DMP will be clearly documented and made accessible to all project members to maintain transparency and accountability. Communication of these updates will be handled proactively to keep partners and stakeholders informed about changes to data management procedures. Periodic compliance checks will also be carried out to verify ongoing adherence to legal and regulatory standards. When updates introduce new practices or modify existing ones, the project will provide targeted training and guidance to ensure personnel are well-prepared to adopt the changes. Ultimately, this review and update process embodies the project's commitment to continuous improvement in data management, with all feedback playing a crucial role in shaping future developments of the DMP.

2.8 Training and Support

To ensure effective data management across the fPVN project, a structured support system will be established, complemented by targeted guidance and resources tailored to the needs of project personnel, collaborators, and external contributors. While introductory training opportunities will be offered, the primary focus will be on ongoing support to assist team members in understanding and applying data management practices in line with the DMP. Support efforts will aim to clarify tasks such as data submission, version control, metadata creation, documentation, and compliance with privacy and security regulations, including GDPR. Resources will be adapted to different roles, such as project partners, data managers, code contributors, and use case owners. A variety of delivery methods will be employed to meet varying preferences, including online tutorials, documentation, one-on-one support, and FAQs. A centralized repository will host all support materials, including best practice guides, templates, checklists, and documentation standards. Role-specific guidance will be made

available to address the unique responsibilities of team members, while periodic updates and refresher materials will help maintain alignment with evolving standards and practices. Feedback mechanisms will be used to continuously improve support offerings, and efforts will be made to ensure accessibility and inclusivity in all materials and communications. English will serve as the primary language for support and documentation, and a simple tracking system will be maintained to monitor engagement with available resources.

3. FAIR Data

3.1 Making Data Accessible According to the FAIR Principles

The AFPVN project is committed to ensuring that all research data generated throughout the project lifecycle adheres to the FAIR principles: Findable, Accessible, Interoperable, and Reusable (Wilkinson et al., 2016). Data will be shared according to the principle "as open as possible, as restricted as necessary," with metadata being made openly available even in cases where data itself cannot be publicly shared due to confidentiality, legal, or ethical constraints.

To enhance discoverability, all archived data will be accompanied by descriptive metadata using standard schemas such as DataCite, ensuring consistency and interoperability. Published datasets, software, and supporting documentation will be assigned persistent identifiers (DOIs), thereby making them easily citable and traceable. Metadata will include relevant keywords and standard descriptors, including project acronym, data type, generation date, responsible partner, and a concise description. Standardized file naming conventions will be established to ensure clarity, uniqueness, and traceability across the project.

The consortium will use open, discipline-specific metadata formats where applicable and ensure compatibility with the European Open Science Cloud (EOSC) standards. Metadata and DOIs will be linked to relevant scientific outputs and project deliverables to ensure a coherent traceability chain from data to publication.

3.2 Making Data Openly Accessible

In line with Horizon Europe's open science policy, the AFPVN project will ensure open access to all peer-reviewed publications, accompanying datasets, software, and relevant metadata needed to validate the research results. In cases where data cannot be fully disclosed as proprietary use case data from industrial partners or data involving sensitive information, partial publication strategies (e.g. metadata-only publication or aggregated formats) will be adopted.

Legal, ethical, or IPR-related restrictions will be evaluated case by case in collaboration with the data-owning partners. In such instances, controlled access mechanisms will be implemented. Where feasible, anonymization or pseudonymization techniques will be applied to maximize openness without compromising personal data protection.

AFPVN-generated code and tools, unless restricted by commercial rights, will also be openly accessible and shared through GitHub and linked to Zenodo for DOI generation. Documentation and training materials relevant to the project's software and data workflows will also be made publicly available through the project website or public repositories.

3.3 Making Data Interoperable

To support interoperability, AFPVN data will follow widely accepted open standards, vocabularies, and formats specific to the data types involved. Metadata will include details on the tools and software used to create and process the data, enabling reuse and reproducibility. Where available, domain-specific ontologies and thesauri will be adopted to enhance semantic interoperability.

3.4 Increasing Data Re-use

The AFPVN project will maximize data reuse by ensuring datasets are well-documented, licensed, and quality-assured. Datasets intended for public release will be published under licenses that support reuse, such as Creative Commons CC BY 4.0 for data and OSI-approved licenses for software. Prior to publication, datasets will undergo internal quality checks to ensure completeness, accuracy, and documentation sufficiency.

To ensure proper timing of data release, datasets associated with publications will be made openly available at the time of publication or as soon as embargo periods (if any) are lifted. When necessary, reuse restrictions due to ethical concerns, industrial confidentiality, or pre-commercial value will be clearly stated in the metadata. The long-term availability of data on open repositories will be planned and discussed with data-hosting platforms to support reuse beyond the lifetime of the project.

4. Ethical Aspects

To the best knowledge of the AFPVN project consortium, there are no anticipated ethical or legal concerns associated with the data generated that would limit its sharing or reuse. All project activities will be carried out in full compliance with the European Code of Conduct for Research Integrity (ALLEA, 2023), the Charter of Fundamental Rights of the European Union (European Parliament, 2000), and relevant EU and national legislation. This includes commitments to safeguard the freedom of research, to protect the physical and moral integrity of individuals, and to respect privacy and data protection principles, particularly in relation to any personal or sensitive data handled during the project. Where human data is involved—such as survey results or user feedback—appropriate consent procedures, anonymization, and data protection measures will be implemented. All research involving human participants, personal data, or ethically sensitive information will undergo internal ethics screening and, where required, review by institutional or national ethics committees. The AFPVN consortium also adheres to the principle of "do no harm", ensuring that all outputs serve public interest without violating ethical norms or partner confidentiality.

5. Conclusions

This document outlines the data involved in the project, detailing its utilization, production, and management during and after the project's completion. It encompasses various data types crucial to the project, including pilot data enriched with open data, project deliverables (both confidential and public), publications, software artifacts, research-related artifacts from the project infrastructure, and working documents.

Furthermore, the deliverable provides insights into methodologies ensuring that research data adheres to the principles of being findable, accessible, interoperable, and re-usable (FAIR). Special attention is given to handling research data throughout and after the project, specifying the data collection, processing, and generation methods, the application of methodologies and standards, considerations for data sharing and open access, and the curation and preservation of data post-project completion.

6. Revision history

6.1 Contributing and reviewing partners

Contributions	Reviews	Participants	Representing partner
1	X	Pär Erik Martinsson	LTU
2	X	Lama Alkhaled	LTU
	x	Jerker Delsing	LTU

6.2 Amendments

No.	Date	Version	Subject of Amendments	Author
1	01-4-2025	0.1	Initial draft	Par Erik Martinsson
2	15-4-2025	0.5	FAIR Principles	Par Erik Martinsson
3	01-05-2025	1	First draft	Par Erik Martinsson

6.3 Quality assurance

No	Date	Version	Approved by
1	29-05-2025	1	Jerker Delsing