D1.4 - Data Management Plan II

Title	D1.4 - Data Management Plan II
Revision Number	3.0
Task reference	T1.2
Lead Beneficiary	UPRC
Responsible	Argyro Mavrogiorgou
Partners	AL, DAEM, GFT, IDSA, INNOV, INNEUROPE, KM, LXS, MOH, NUIG, UBI, UNP
Deliverable Type	Report
Dissemination Level	PU
Due Date	2025-12-31 [Month 36]
Delivered Date	2025-12-19
Internal Reviewers	JOT GFT
Quality Assurance	UPRC
Acceptance	Coordinator Accepted
Project Title	FAME - Federated decentralized trusted dAta Marketplace for Embedded finance
Grant Agreement No.	101092639
EC Project Officer	Stefano Bertolo
Programme	HORIZON-CL4-2022-DATA-01-04

This project has received funding from the European Union's Horizon research and innovation programme under Grant Agreement no 101092639

Revision History

Version	Date	Partners	Description
0.1	2025-08-20	UPRC	ToC
0.2	2025-10-02	UPRC	Initial content
0.3	2025-12-01	UPRC	Integration of information from partners
1.0	2025-12-05	JOT, GFT	Submission for Internal Review
1.1	2025-12-12	UPRC	Reviewed version
2.0	2025-12-14	UPRC	Final version ready for QA
2.1	2025-12-16	GFT	Final version reviewed by QA
3.0	2025-12-19	UPRC	Final version ready for submission

Definitions

Acronym	Definition
4AML	4th Anti Money Laundering Directive
AAI	authentication authorization infrastructure
AI	Artificial Intelligence
API	Application Programming Interface
ATOS	Atos It Solutions And Services Iberia Sl
AWS	Amazon Web Services
CDS	Copernicus Data Store
CDTI	Technological Development and Innovation Centre (from Spanish “Centro para el Desarrollo Tecnológico y la Innovación)
CSS	Cascading Style Sheets
CSV	Comma Separated Value files
DAEM	Dimos Athinaion Epicheirisi Michanografisis
DB	Data Base
DCAT	Data Catalog Vocabulary
DOI	Digital object identifier
DPO	Data Protection Officer
EC	European Commission
ESG	Environmental, Social and Governance
EU	European Union
FAIR	Findable Accessible Interoperable Reusable
FAME	Federated decentralized trusted dAta Marketplace for Embedded finance
FDAC	Federated Data Assets Catalogue
FIBO	Financial Industry Business Ontology
FIGI	Financial Instrument Global Identifier
FML	Federated Machine Learning
GA	Grant Agreement|General Assemply
GDPR	General Data Protection Regulation
HTML	Hypertext Markup Language
HTTP	HyperText Transfer Protocol
ID	Identity
IDSA	International Data Spaces Association
IEC	International Electrotechnical Commission
IEEE	Institute (of) Electrical (and) Electronic Engineers
INNOV	Innov-Acts Limited
IP	Internet Protocol

IPR	Intellectual Property Right
ISO	International Organization for Standardization
JDBC	Java Database Connectivity
JRC	Jrc Capital Management Consultancy and Research Gmbh
JSI	Institut Jozef Stefan
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LLM	Large language model
LSTM	Long Short-Term Memory
ML	Machine Learning
MOH	Motor Oil (Hellas) Diilistiria Korinthou A.E.
MPI	Message Passing Interface
NLP	Natural language processing
NOVO	Novomatix Idiotiki Kefalaiouchiki Etaireia
NRS	Norsk Regnesentral
ORM	Object-Relational Mapping
PAT	Pricing Advisory Tool
PC	Personal Computer
PDF	Portable Description Format
PSD2	2nd Payment Services Directive
RBAC	Role-Based Access Control
RDF	Resource Description Framework
REST	Representational State Transfer
RFM	Recency, Frequency, and Monetary
SA	Solution Architecture
SCSS	Source Code Control System
SDK	Software Development Kit
SHA	Secure Hash Algorithm
SHAP	SHapley Additive exPlanations
SMOTE	Synthetic Minority Over-sampling Technique
SQL	Structured Query Language
TFEU	Treaty on the Functioning of the European Union
TLS	Transport Layer Security
TM	Trade and Monetisation
TXT	Text|TeXT file
UI	User Interface

URL	Uniform Resource Locator
UUID	Universally Unique Identifier
VDIH	Virtual Digital Innovation Hub
WP	Workpackage
XAI	Explainable Artificial Intelligence
XML	Extensible Markup Language

Other acronyms and abbreviations not present in the table, are introduced in the text along with their definitions.

Executive Summary

This deliverable presents the final Data Management Plan (DMP) of the FAME project and confirms that, at the conclusion of its 36-month lifecycle, FAME has established a robust, auditable, and sustainable framework for managing all data collected, processed, and generated under Horizon Europe.

The DMP fully complies with the European Commission's Horizon Europe Data Management Plan template and systematically applies the FAIR principles to ensure that FAME data remains findable, accessible, interoperable, and reusable, while respecting GDPR, intellectual property rights, and contractual obligations. The plan reflects the final operational state of data management across all Work Packages and consortium partners.

Over the project lifecycle, FAME generated and managed more than 3,600 documents, covering administrative coordination, technical development, scientific output, communication, and evaluation activities. These data assets include internal project data, software and solution prototypes, pilot and evaluation datasets, and externally exploitable data assets indexed within the FAME Data Marketplace. The DMP defines consistent procedures for documentation, metadata, formats, access rights, licensing, security, and long-term preservation across this full data landscape.

To support open science and long-term impact, FAME systematically used trusted repositories. In particular, 52 validated project outputs have been published on Zenodo, ensuring persistent identifiers, standardized metadata, and continued discoverability and accessibility beyond the project's end. This guarantees that key FAME results remain reusable by the research community, policymakers, and industry stakeholders after project completion.

FAME operationalizes FAIR principles through concrete and verifiable implementation measures, including standardized metadata schemas, persistent identifiers, role-based access control, and clear licensing models. Where data cannot be openly shared due to privacy, confidentiality, or commercial sensitivity, the DMP ensures that metadata remains accessible, preserving transparency and enabling future controlled reuse in line with Commission guidance.

This final DMP incorporates insights from all project tasks and deliverables and reflects lessons learned throughout the project lifecycle. It confirms that FAME has minimized data-related risk, ensured regulatory and ethical compliance, and maximized the scientific, technical, and economic value of its data assets.

FAME therefore concludes with a FAIR-aligned, well-governed, and sustainable data foundation that directly supports Horizon Europe objectives and enables future innovation and reuse beyond the lifetime of the project.

Table of Contents

1 Introduction 6
1.1 Objective of the Deliverable 7
1.2 Insights from other Tasks and Deliverables 7
1.3 Structure 7
1.4 Updates from previous version 7
2 Methodology 9
2.1 FAIR management of research data 9
2.2 High-quality Data Generation Best Practices 10
2.2.1 Documentation 11
2.2.2 Metadata 11
2.2.3 Data Formats 12
2.2.4 Access to Data 13
2.2.5 Persistent Identifiers 13
2.2.6 Data Licenses 14
2.3 Collection of information for FAME Data Management Plan 14
3 FAME Data Summary 18
3.1 Types of Collected and Generated Data 18
3.2 Formats of the Data 19
3.3 Origin of the data 20
3.4 Size and Volume of the Data 22
3.5 Data Utility 22
3.6 FAME Data 22
3.6.1 FAME Internal Data 22
3.6.2 FAME External Data 29
4 FAIR Data & Zenodo Platform in FAME 121
4.1 FAIR Data 121
4.1.1 Making data findable, including provision for metadata 122
4.1.2 Making data accessible 123
4.1.3 Making data interoperable 123
4.1.4 Increase data re-use 124
4.2 FAME FAIR Data through the Zenodo Platform 124
4.2.1 Zenodo Platform in a Nutshell 124
4.2.2 Zenodo Platform in FAME 124
4.2.3 Zenodo throughout FAME lifecycle 125

5 Allocation of Resources 126
5.1 Costs for Making Data FAIR 126
5.2 Data Management Responsibilities 126
6 Data Security 127
7 Ethical Aspects 129
8 Conclusions 130
References 131
Annex I 132
Annex II 133

List of Figures

Figure 1 - FAME on Zenodo 125

List of Tables

Table 1 – FAIR guiding principles 9
Table 2 – Data collected and/or generated in FAME 18
Table 3 – Data formats of the FAME data 19
Table 4 – FAME Marketplaces and Data Spaces 20
Table 5 – Internal communication data 22
Table 6 – Partners' data 23
Table 7 – Internal administrative data 24
Table 8 – Deliverables 24
Table 9 – Publications 25
Table 10 – Users' registration data 25
Table 11 – Solution prototypes 26
Table 12 – Software code 26
Table 13 – Co-design activities data 27
Table 14 – Evaluation data 28
Table 15 – Recommendations' data 28
Table 16 – Park & Ride Service 29
Table 17 – Loyalty Program Service 30
Table 18 – Aggressive ESG-Selective Portfolio 31
Table 19 – Aggressive Growth Portfolio 33
Table 20 – Aggressive Sustainable Portfolio 34
Table 21 – Balanced ESG-Aware Portfolio 35
Table 22 – Balanced Impact Portfolio 36
Table 23 – Balanced Standard Portfolio 38
Table 24 – Conservative Conscious Portfolio 39
Table 25 – Conservative Sustainable Portfolio 40
Table 26 – Conservative Traditional Portfolio 41
Table 27 – Ranking Aggressive ESG-Selective Portfolio 42
Table 28 – Ranking Aggressive Growth Portfolio 43

Table 29 – Ranking Aggressive Sustainable Portfolio 45
Table 30 – Ranking Balanced ESG-Aware Portfolio 46
Table 31 – Ranking Balanced Impact Portfolio 47
Table 32 – Ranking Balanced Standard Portfolio 48
Table 33 – Ranking Conservative Conscious Portfolio 49
Table 34 – Ranking Conservative Sustainable Portfolio 50
Table 35 – Ranking Conservative Traditional Portfolio 52
Table 36 – Service Asset: Pilot 5 ESG Ranking & Portfolio Service (PaaS) 53
Table 37 – Extreme heat projections, CCCma-CanESM2/CCCma-CanRCM4 54
Table 38 – Extreme heat projections, CCCma-CanESM2/ OURANOS-CRCM5 56
Table 39 – Extreme heat projections, CNRM-CERFACS-CM5/ OURANOS-CRCM5 57
Table 40 – Extreme heat projections, MPI-M-MPI-ESM-LR/ OURANOS-CRCM5 59
Table 41 – Extreme heat projections, NOAA-GFDL-ESM2M/OURANOS-CRCM5 61
Table 42 – Storm loss projections, CNRM-CERFACS-CM5/CNRM-ALADIN63 63
Table 43 – Storm loss projections, ICHEC-EC-EARTH/DMI-HIRHAM5 64
Table 44 – Storm loss projections, MPI-M-MPI-ESM-LR/SMHI-RCA4 66
Table 45 – MOH Compressor Sensor Datasets 68
Table 46 – Pricing Advisory Tool 69
Table 47 – IIoT Data Quality Assessment Service 72
Table 48 – CO2 Emissions Monitoring API 74
Table 49 – StreamStory Causal (SSCausal) 75
Table 50 – StreamStory XAI Dashboard 77
Table 51 – StreamStory Python Client 78
Table 52 – Incremental Analytics 79
Table 53 – Fast Data Importer 80
Table 54 – FML Framework 82
Table 55 – Marketplace UI Template 83
Table 56 – Marketplace Chatbot 85
Table 57 – FAME Marketplace Information Package 87
Table 58 – Anomaly Detection FL Center Server 88
Table 59 – Anomaly Detection FL Client 90
Table 60 – Anomaly Detection FL Root Server 91
Table 61 – Federated KMeans Server for Financial Services 93
Table 62 – Federated KMeans Client for Financial Services 94
Table 63 – SD API-Server 96
Table 64 – SD Node-Resolver 97
Table 65 – SD Model-Server 98
Table 66 – SD WebApp 100
Table 67 – LSTM Forecasting 101
Table 68 – TabNet Classifier Model 103
Table 69 – K-Means Model 104
Table 70 – K-Means Model 106
Table 71 – K-Means Model 107
Table 72 – LSTM Model 109
Table 73 – RFM Processing Pipeline 110
Table 74 – RFM Processing Pipeline 111
Table 75 – Parking Zones Dataset 113
Table 76 – RFM Dataset Pilot 2 115

Table 77 – RFM Dataset Pilot 1 116
Table 78 – Pilot 4 Synthetic Dataset 118
Table 79 – Sensor Anomaly Detector 119
Table 80 – FAME FAIR Sub-Principles 122

1 Introduction

A Data Management Plan (DMP) in European Union (EU) projects is developed to describe how the data will be gathered, processed and/or generated, as well as how it will be stored. It is an essential component of effective data management as it systematically describes how datasets will be used by project partners and what data assets will be generated within the scope of the project. This will encourage the project partners to make data management decisions based on FAME project and FAIR principles [2], while also facilitating the smooth operation of the project.

FAME project develops a secure federated data marketplace for embedded finance (EmFi), aiming to demonstrate the full potential of the data economy, and in this context the current deliverable focuses on the management of the data collected, processed, or generated within the project and its lifecycle by all its consortium partners. The DMP aims to monitor the generated data regarding their privacy and confidentiality, ensuring that the legal and ethical standards for data generation, use, storage, and sharing are met throughout the project lifecycle and in line with the overall management of FAME, as also foreseen in the Grant Agreement (GA) and Consortium Agreement (CA).

It should be made clear that when speaking about data in a DMP, this does not refer solely to datasets produced or used in the scope of the project's scientific and technological developments, but rather to any data assets produced and managed by the project. More specifically, it may refer to:

- *Internal administrative data*, such as: (i) Data from administrative and financial management (e.g., partner contact information), (ii) Meeting/Web conference related material (e.g., participant's list, agenda, meeting minutes kept during a conference call or a general assembly meeting), (iii) Templates (e.g., deliverables' templates, presentations' templates), as well as (iv) Documentation of communication among members of the project (e-mails, instant messaging) (e.g., e-mails exchanged internally for FAME communication purposes)
- *External communication data*, such as: (i) Deliverables and reports (e.g., FAME deliverables), (ii) Publications (e.g., conference papers, journal articles), (iii) Data from the marketing/dissemination/communication/commercialization process (e.g., blog posts, social media posts, photographs from participation to events)
- Technical data, namely: (i) Software code (e.g., source code), and (ii) Prototype (e.g., solution prototypes)
- *Proof-of-concept data*, such as: (i) Datasets (e.g., open datasets, proprietary datasets, visualizations)
- *Data assets*, namely assets indexed in the FAME Data Marketplace, following the path defined through the technical implementations of WP2, WP3, WP4, and WP5 of the project

Information to produce this deliverable was gathered by the FAME Consortium through two (2) questionnaires (ANNEX I and ANNEX II) with a variety of questions per data management aspect and in respect to the FAIR data principles, targeting all the project partners' data (ANNEX I), as well as specifically the data assets indexed in the FAME Data Marketplace (ANNEX II) through investigating more detailed information per data asset (e.g., policies for re-use, archiving and providing access, responsibilities). The structure of the DMP is also based on European Commission's (EC) DMP template. The DMP reflects a current picture of the Project based on the answers from each partner and was viewed as a living document that was supplemented and refined as new issues and information came to light during the project's 36-month lifecycle.

1.1 Objective of the Deliverable

FAME's DMP, aims to identify and organize all data collected, processed or generated throughout the project's lifecycle, as well as specific standards used, data management policies and potential best practices showcased by the consortium partners.

1.2 Insights from other Tasks and Deliverables

D1.4 is inherently connected with all FAME Tasks and Work Packages (WP) since its content is tightly coupled with all the project activities, as it is described in Section 3. During the lifecycle of each task, data and data assets are being generated that are properly managed through the project's DMP, following all related guidelines, principles, best practices and FAIR strategies.

1.3 Structure

The rest of the deliverable is structured as follows:

- **Section 2** provides the methodology of this work and how the FAIR principles are applied in the project, including explanations of the basic terminologies that are used throughout the current deliverable
- **Section 3** provides an overview of the FAME data, providing additional information on the types and content of the collected and generated data
- **Section 4** focuses on the FAIR principles in FAME, including the techniques that are used for making FAIR data, as well as a deeper explanation of the FAME Zenodo platform
- **Section 5** briefly replies to the questions related to the allocation of resources
- **Section 6** and **Section 7** cover all the ethical and security aspects
- **Section 8** concludes the current deliverable, introducing our lessons learned

1.4 Updates from previous version

The current deliverable (*D1.4 – Data Management Plan II*), includes the below updates with regards to its previous version (*D1.2 – Data Management Plan I*) [3]:

- **Executive Summary**

The text has been enhanced, specifically referring to the current updates of the current deliverable

- **Section 1 - Introduction**

The Structure of the deliverable has been adapted, to reflect its current positioning

- **Section 2 – Methodology and Basic Terminology**

The **High-Quality Data Generation Best Practices (Section 2.2)** have been adapted accordingly, with additional explanatory text for each of its subsections, including multiple examples on the approaches that are being followed in the context of the FAME project

The **Collection of Information for FAME DMP (Section 2.3)** sub-section has been enhanced explaining in deep detail the content of the questionnaires that were provided to the project partners regarding the DMP, tailored to gather additional input for their data and data assets already indexed to the FAME Data Marketplace

- **Section 3 – FAME Data Summary**

The **Types of Collected and Generated Data (Section 3.1)** has been restructured

The **Formats of the Data (Section 3.2)** have been updated accordingly, based on the latest input of the FAME partners to cover all the available proof-of-concept data

In relation to the **FAME data (Section 3.6)**, new tables have been provided for the FAME Internal and FAME External data covering all the new topics that the FAME Data Marketplace has introduced (in relation to the FAME data assets)

- **Section 4 – FAIR Data & Zenodo Platform**

The whole Section has been restructured in order not only to refer to the techniques that FAME is following to making FAIR data, but also to explicitly refer to the way that the Zenodo platform is exploited for achieving the abovementioned data characteristics and overall openness. For that reason, a new sub-section namely "**Section 4.2 - FAME FAIR Data through the Zenodo Platform**" has been created, including all the relevant information and procedures followed to make FAME data assets available on the Zenodo platform

- **Section 5 – Allocation of Resources** & **Section 6 – Data Security**

In both Sections no major updates have been provided, apart from some text refinements and additions

- **Section 7 – Ethical Aspects**

The overall Section has been enhanced, speaking specifically for the steps that have been followed internally in relation to the Ethical Board (EB), and FAME's further Ethical Aspects-related steps

- **Section 8 – Conclusions**

This is a new section, where the lessons learned from the 36 months of the project have been included, as well as our next steps, after the project's end

- **References**

The already existing references have been checked, whereas new references have been added citing the newly inserted/upated text in the above Sections

2 Methodology

2.1 FAIR management of research data

One difficulty that data-driven projects like FAME face is making it easier for people to share expertise. To improve the FAIR paradigm of digital assets, the FAIR Principles for Scientific Data Management and Stewardship [4] were released in 2016. In this part, the recommended practices for making data FAIR in FAME are introduced. The FAIR principles, among other things, offer a continuum of growing reusability for data and digital objects across a variety of implementations (including data-related algorithms, tools, workflows, protocols, services, and other types of digital and research items). These solutions outline the capabilities and features of systems and services that can be used to produce beneficial research outputs, have them reviewed, and have them widely repurposed while giving the creators and users proper credit. Well-described, accessible, and standard data are essential for finding relevant data, performing machine analysis, and employing next generation techniques (e.g., Artificial Intelligence (AI), eXplanaible AI (XAI), Machine Learning (ML)).

The FAIR principles have gained a lot of support from EU authorities as a quality standard and sparked discussions on data stewardship in open science and data-driven research around the world [5]. They have moreover forced financing organizations to talk about how they will carry out their plans. Some of these needs are currently being established, but others already have a fully developed set of rules. This strategy typically comes before implementation decisions and does not imply a specific set of technical specifications, guidelines, or solutions. However, they place a focus on "machine actionability" which is the capacity of a computer system to locate, access, interact with, and reuse data without much help from humans. A summary of the FAIR guiding principles can be found in Table 1:

Table 1 – FAIR guiding principles

Findable	• Meta(data) are assigned a globally unique and persistent identifier • Data are described with rich metadata • Metadata clearly include the identifier of the data it describes • Meta(data) are registered or indexed in a searchable source
Accessible	• Meta(data) are retrievable by their identifier using a standardized communication protocol • The protocol is open, free, and universally implementable • The protocol allows for an authentication and authorization procedure, where necessary • Metadata are accessible, even when the data are no longer available
Interoperable	• Meta(data) uses a formal, accessible, shared and broadly applicable language for knowledge representation • Meta(data) use vocabularies that follow FAIR principles • Meta(data) includes qualified references to other meta(data)
Reusable	• Meta(data) are richly described with a plurality of accurate and relevant attributes • Meta(data) are released with a clear and accessible data usage license • Meta(data) are associated with detailed provenance

- Meta(data) meet domain-relevant community standards

Simply expressed, the **Findability** principle dictates that data must be clearly and unambiguously identified, described, registered, or indexed. The primary properties of the data should be methodically documented, preferably in a standard format, and they should be indexed and stored in a public repository, such as a data archive. The data should be also given a distinctive and persistent identification. Data should be accessible through a clearly defined access mechanism, preferably through automated means, according to the principle of **Accessibility**. It entails setting up authentication and authorization processes for access as well as, if necessary, automated protocols for data retrieval. Data and metadata must be conceptualized, communicated, and organized in accordance with widely accepted, publicly accessible standards based on the **Interoperability** principle. Standard data formats, variables, ontologies, and other techniques shall be used. The idea of **Reusability** states that data features and provenance should be thoroughly specified in accordance with community standards relevant to the domain, and that usage instructions should be simple to understand. Accurate and pertinent data descriptions, access and use permissions, community standards, and provenance for each data item are all required as part of this process.

Data life cycle management should use these four (4) principles, which are closely interrelated. The open sharing of data is not always required when the FAIR principles are applied to the workflow of a research effort. In this context, it should be made clear that open data and FAIR data are not the same thing. Open data is “information that is freely available for anyone to access, use, and share for any purpose”. However, the FAIR principles offer “recommendations for sharing data while also respecting any potential legal, moral, and contractual limitations” [6].

2.2 High-quality Data Generation Best Practices

The following best practices are intended to assist project partners in generating high-quality data, which will enhance research production and impact. It will give the data more structure so that other researchers can utilize it again. Additionally, it will enhance peer review and provide value to research, maintaining high levels of research integrity. Applying the FAIR data principles will therefore increase FAME’s transparency and reproducibility. FAIR principles are applied differently depending on the discipline and research method. However, there are several common elements to consider within search workflows such as:

- **Documentation**: By giving context, good documentation will guarantee that the data are understood by others.
- **Metadata**: When data is accompanied by metadata, it is easier to find it.
- **Data formats**: When integrating data or converting data to a machine-readable format, several formats must be considered.
- **Access to data**: The user controls who has access to which data and under what circumstances.
- **Persistent identifiers**: Through the creation of a link that remains with the data for a long time, persistent identifiers make it possible for others to locate and use the data.
- **Data licenses**: The legal contract defining how other parties may use the data is known as a license.

All in all, persistent identifiers and metadata are required to make data Findable. For data to be Accessible, it is important to specify who has access to the data and how. If it is not possible to publish data publicly due to data protection and IPR constraints, it is advised to enable access to

metadata, for instance through a data repository. To promote Interoperability, common standards and open data formats should be employed. Finally, the data can be Reused in a variety of ways and will be easier for others to understand if suitable documentation is added to them.

The latter is also in line with Directive (EU) 2019/1024 [7] on open data and re-use of public sector information (Open Data Directive) and Implementing Regulation (EU) 2023/138 [8] laying down a list of specific high-value data and their arrangements for their publication and re-use and in particular, the applicable conditions for re-use and the minimum requirements for disseminating data via Application Programming Interfaces (APIs).

Following, a description of the best practices for generating data of high-quality is being provided.

2.2.1 Documentation

Documentation within the FAME project is crucial and often complex, due to the heterogeneity, volume, and lifecycle diversity of data assets exchanged through a federated data marketplace. As FAME aims to enable trusted sharing, discovery, and exploitation of data assets across multiple stakeholders and domains, systematic documentation is essential to ensure transparency, traceability, interoperability, and compliance with regulatory and ethical requirements. From the beginning of the project, it is crucial to agree on a commonly accepted methodology for organizing and documenting the data, with the active contribution of all stakeholders. Each data asset should made available through the FAME Data Marketplace is uniquely identified, accompanied by complete metadata, and clearly documented as to its origin, collection methods, purpose, application, and licensing constraints.

In the context of FAME, documentation may include datasets originating from industrial systems, digital platforms, sensors, enterprise information systems, or derived data products created through analytics or AI-based processing pipelines. Records associated with these data assets document, where applicable: (i) when, how, and by whom the data was generated or provided, (ii) the conditions under which it can be accessed and reused, (iii) the versions of the data asset and related metadata, (iv) any transformation, anonymisation, or enrichment processes applied prior to publication in the marketplace, and (v) usage, consumption, or transaction outcomes within the platform. The use of appropriate tools for metadata management, Data Version Control (DVC) for data and model version control, provenance tracking, as well as standards such as FAIR and ISO/IEC 25012 for data quality, can significantly enhance the quality of documentation. Finally, a well-organized and transparent documentation of the research process allows for the reproducibility of results, compliance with ethical and legal standards, and the facilitation of future exploitation of data by both stakeholders and the wider research community.

2.2.2 Metadata

An equally critical type of documentation in the context of FAME is metadata, that is, information that describes, interprets, or accompanies data assets published, exchanged, or consumed via the FAME Data Marketplace. Metadata makes it possible to correctly classify, understand, connect, and reuse data. It includes information such as location, date and time of collection, file size and format, technical characteristics, keywords, but also contextual information (e.g., keywords, domain vocabularies, intended purpose, ownership, responsible entity, usage conditions, references to relevant ontologies or data models).

Within FAME, where data assets may originate from heterogeneous sources and are shared among multiple stakeholders, metadata plays a central role in enabling interoperability, automation, and trust. For example, a data asset offered through the FAME marketplace for operational optimisation or decision support may require metadata describing the data source, temporal and spatial scope, quality characteristics, update frequency, applicable constraints, and assumptions under which the data was generated. Such metadata enables data consumers to correctly assess the fitness-for-purpose of the asset, supports automated discovery and matchmaking within the marketplace, and facilitates the transfer or reuse of data across comparable use cases and domains.

The existence of metadata standards (such as Dublin Core, schema.org, ISO/IEC 11179 or, for more technical applications, OpenMINDS or OPC UA standards for industrial data) is crucial to ensure consistency and automatic processing. However, in cases where historical or unstructured data is used, metadata may be absent or not harmonized with a standard. In this case, it is possible to create synthetic metadata through AI tools - e.g., using Natural Language Processing (NLP) to automatically extract descriptive features or classification/clustering methods for ontological grouping.

It is important to note that when preparing datasets for training AI models, a significant amount of time is spent interpreting and recording metadata, especially when the data comes from different subsystems. The use of standardized and well-designed metadata significantly reduces this time and enhances Findability, Interoperability and Reusability - key principles of FAIR data.

As part of the FAME project, partners are expected to adopt metadata best practices and to provide adequate documentation for each data asset made available through the marketplace, even in the absence of universally adopted standards. At a minimum, this may include structured descriptions and human-readable documentation (e.g., README files) detailing the content of the data asset, its structure, variables, units of measurement, quality indicators, constraints, and usage conditions. The consistent application of these practices strengthens trust among marketplace participants, improves the reliability and transparency of data exchanges, and supports the sustainable exploitation of data assets within and beyond the lifetime of the FAME project.

2.2.3 Data Formats

Data formats are a critical factor for the effective collection, sharing and exploitation of data. When data assets are made available using widely adopted, well-documented, and open formats, their interoperability and reusability are significantly enhanced, both for human users and for automated systems operating within data-driven platforms. Conversely, the use of undocumented, proprietary, or closed binary formats can hinder data inspection, reproducibility, and integration, thereby limiting the value and uptake of data assets within a data marketplace environment.

In the context of FAME, where data assets originate from heterogeneous sources and are exchanged across organisational and technical boundaries, the selection of appropriate standardised formats (e.g., CSV, JSON, XML, HDF5, Parquet, or equivalent open formats) is essential. The use of such formats directly supports the FAIR data principles and facilitates the seamless integration of data assets into analytics pipelines, AI-based services, and simulation or decision-support frameworks. Moreover, reliance on open and standard formats contributes to regulatory compliance, transparency, and the responsible and ethical use of data, particularly in scenarios involving automated processing or AI-enabled services.

To ensure maximum reusability and interoperability, it is advisable to accurately document the format of the files, declare the structure of the data (e.g. headers, data types, encoding), and, where possible, provide an accompanying example or schema that facilitates machine reading. An example of a good practice is to publish a dataset in CSV format along with an accompanying YAML/JSON Schema file that defines the fields and validation rules.

Overall, the strategic selection and documentation of data formats is a foundational element in ensuring that data assets exchanged through the FAME Data Marketplace can be efficiently discovered, accessed, integrated, reproduced, and exploited across a diverse and interdisciplinary ecosystem of data providers, consumers, and service developers, both during and beyond the lifetime of the project.

2.2.4 Access to Data

In cases where data protection restrictions are applicable, it is considered critical to document access to data. Even if the data itself cannot be shared or published freely, the metadata should remain accessible and well documented, to inform the interested party about how and under what conditions the full data can be accessed. For example, in scenarios where data assets include personal, sensitive, or commercially confidential information, the raw data may be subject to strict access limitations. Nevertheless, metadata can include structured information such as: (i) whether the data is accessible or restricted, (ii) the entity or controller that holds the rights, (iii) the legal framework under which it was collected (e.g. consent, contract, public interest), (iv) the access request process, including a contact email or link to a relevant form, or (v) the licenses.

Embedding such access-related information within metadata enhances transparency, supports compliance with data protection and governance requirements (including General Data Protection Regulation (GDPR) and data sovereignty principles), and reinforces the ethical and responsible use of data within the FAME ecosystem. In addition, the use of role-based or policy-based access control mechanisms can be explicitly described and referenced through metadata, enabling both human users and automated systems to determine in advance whether a given data asset can be accessed or reused and under which conditions. The presence of such fields in metadata not only enhances transparency but also reduces time and uncertainty for those who wish to analyze, reuse, or extend the data - while promoting Responsible Research & Innovation (RRI) promoted under Horizon Europe.

2.2.5 Persistent Identifiers

To ensure that research data is easily located and accessible, it is necessary to assign them a permanent and unique identifier. Such identifiers allow for consistent reference to a specific dataset or digital object, regardless of its location or hosting system. An example can be the use of Digital Object Identifiers (DOIs), which are widely used in scientific publications to accurately reference articles, datasets or software. A DOI leads the user to a fixed URL, through which they can access the corresponding content, under the terms defined by the publisher or creator.

In the context of FAME, where multiple versions of data assets may be published, updated, exchanged, or reused through the data marketplace, the assignment of persistent identifiers plays a key role in version control, traceability, and lifecycle management. For example, a curated data asset made available through the FAME marketplace for analytics or service development purposes may be registered in a trusted repository or catalogue and assigned a persistent identifier (such as a DOI or equivalent). Even if the technical hosting location changes, or access is mediated through

different marketplace components, the identifier continues to resolve to the authoritative and archived representation of the data asset and its metadata.

Using persistent identifiers not only facilitates access, but also enhances transparency, traceability, and reproducibility of results. Within the framework of the FAIR principles, the existence of a stable identifier is a prerequisite for the inclusion of data in public databases or their further exploitation through automatic systems. Finally, it is good to accompany the identifier with relevant documentation (metadata) describing the content, purpose, language, license and publication date. A well-designed data identification system improves collaboration among partners, speeds up the research process and minimizes the risk of data misuse.

2.2.6 Data Licenses

When making data assets openly available or shared through the FAME Data Marketplace, it is essential that they are accompanied by a clear data license that specifies how and by whom the data can be used, ensuring legal clarity and fostering responsible data reuse. The most common types of such licenses are Creative Commons (CC), which provide flexible licensing models, from completely open to restrictive schemes. This form of license is particularly suitable for projects such as FAME, where interoperability and data reuse by other research groups is critical.

Within FAME, where data assets may be contributed by multiple stakeholders, the selection of an appropriate licence is a critical governance decision. Depending on the nature of the data asset, its sensitivity, and its intended exploitation model, this may include standard Creative Commons licences (e.g., CC BY, CC BY-SA), open data licences such as Open Data Commons Attribution (ODC-BY), or, where applicable, specific licensing schemes for software components or APIs (e.g., MIT, Apache 2.0). In certain cases, custom or contractual licences may also be applied to reflect commercial, confidentiality, or sector-specific constraints while still enabling controlled access and reuse through the marketplace.

For example, a data asset published via the FAME marketplace for analytics, benchmarking, or AI-based service development may be released under a CC BY 4.0 licence, allowing reuse and adaptation by third parties subject to appropriate attribution. Clear data licensing not only protects creators and guides users but is also a prerequisite for the full implementation of the FAIR principles, especially Reusability. Finally, the metadata should include the selected license with a link to the full legal text, to facilitate automatic processing by AI systems and search engines.

2.3 Collection of information for FAME Data Management Plan

The information for this deliverable was gathered by the FAME consortium through two (2) questionnaires dedicated to the FAME partners in general (FAME Internal Data - ANNEX I), and to the FAME partners that have indexed a list of proprietary data assets to the FAME Data Marketplace at the current stage of the deliverable (FAME External Data - ANNEX II) (considering both technical and pilot partners). It is based on the “Guidelines on FAIR Data Management” [4] and corresponds to an updated template associated with them (Horizon Europe DMP Template). The leaders of this deliverable gave the partners the final version of the questionnaires after making several changes. The questionnaire that was dedicated to the FAME partners in general (ANNEX I), required information related to the below:

- **Data description**
  - It may include a short description of the nature and type of data that will be collected or generated

- **Related WP/Task**
  - It specifies the WP and Task(s) where the data is created or used, facilitating into tracing data back to the research context and responsible partners
- **Data origin**
  - It indicates how the data is obtained, referring to whether the data is generated in-house, collected via third-party systems, or sourced from existing datasets
- **Will you re-use any existing data? If yes, how?**
  - It clarifies if open or proprietary datasets will be reused, including the source, relevance, and method of integration
- **Methodologies for data collection/generation**
  - It describes the methodologies used to collect or generate the data, including protocols for ensuring accuracy and consistency
- **Data format**
  - It outlines the file types and encoding standards used, ensuring data is in a machine-readable and interoperable format
- **Where this data will be stored?**
  - It details the storage infrastructure, along with backup and access provisions
- **Expected size of the data**
  - It provides an estimation of the total data volume, helping with infrastructure requirements
- **Metadata and standards used**
  - It lists the metadata schema and the applied documentation standards, helping to describe and contextualize the data for its stakeholders
- **For whom might the data be useful?**
  - It identifies the target users of the data beyond the project partners, such as researchers, policymakers, or the wider AI community
- **Data access, sharing and licensing**
  - It explains how and when the data will be shared, through what platforms, and under which license

In addition, the questionnaire that was dedicated to the FAME owners who had already indexed their data assets to the FAME Data Marketplace (ANNEX II) was split in specific categories, requiring for more detailed and specific information in relation to the: (i) Data Asset Overview, (ii) Metadata, (iii) Storage, backup & Security, (iv) Protection & Privacy, (v) Access & Sharing, (vi) Archiving & Providing Access, and (vii) Responsibility. A description of the required information is provided below:

Data Asset Overview

- **Description**
  - It deals with an overview of the data asset's contents, structure, variables and intended use
- **Source/origin**
  - It provides the original source of the data asset or links to external repositories or documentation
- **Collection/Generation methodology**
  - It outlines the scientific or procedural methodologies used to collect/generate the data asset
- **Type**
  - It describes the type of the data asset
- **Format**
  - It lists the data formats used to store the data asset

- **Behavior**
  - It indicates whether the data asset is fixed or dynamic and the way that it will be updated
- **Expected size**
  - It provides an estimation of storage needs, to support infrastructure planning

Metadata

- **Metadata & standards used**
  - It specifies the metadata schema and standards used for the data asset
- **Documentation methodology**
  - It explains how the data asset will be documented and the methodology that this will be achieved

Storage, Backup & Security

- **Storage methodology**
  - It specifies the storage methodology of the data asset, as well as its access permissions
- **Backup methodology**
  - It describes the backup frequency, location, and tools used to ensure the data asset's recovery
- **Security methodology**
  - It includes security measures like encryption, access control, audit trails, and secure transmission protocols to prevent unauthorized data access

Protection & Privacy

- **Privacy & confidentiality issues**
  - It explores how the project addresses privacy and confidentiality rights for the data asset
- **Intellectual property rights responsible**
  - It identifies who holds the IPR and what agreements may define ownership
- **Intellectual property rights constraints**
  - It clarifies whether any restrictions exist on using, sharing, or publishing the data asset

Access & Sharing

- **Receiver(s)**
  - It indicates the potential end-users, considering the internal and external research community
- **Re-use methodology**
  - It explains how external stakeholders can access, interpret, and apply the data asset
- **Any other existing data (asset) re-use**
  - It details sources of pre-existing data or data assets and the way that they will be used
- **Access, sharing & licensing methodology**
  - It outlines platforms, licenses, and criteria for granting access

Archiving & Providing Access

- **Preservation planning**
  - It outlines how the data asset will be preserved beyond the project's duration
- **Long-term access provision planning**
  - It ensures the planning for future users to be able to locate and access the data
- **Preservation duration**
  - It provides the data retention period

Responsibility

- **Responsible for data asset management**
  - It designates the responsible for implementing the data asset's management,

	including full communication information (i.e., Name Surname; Organization name; Email address)

3 FAME Data Summary

The following Sections provide information on the data that has been collected and generated based on the project's activities.

3.1 Types of Collected and Generated Data

In FAME both quantitative (i.e., data in the form of numbers) and qualitative (i.e., non-numerical) data are being collected. The identified data can be classified into ***FAME Internal Data*** (i.e., *data collected/generated for internal project administration, implementation, communication & evaluation purposes*), and ***FAME External Data*** (i.e., *data assets provided to the external world, indexed within the FAME Data Marketplace for further external usage*) according to the following typology:

- **FAME Internal Data**
  - Data collected/generated through internal administration procedures
  - Data collected/generated for communication procedures
  - Data collected/generated for technical purposes
  - Data collected/generated for FAME evaluation
- **FAME External Data**
  - Data (currently) indexed within the FAME Data Marketplace from the FAME pilots
  - Data (currently) indexed within the FAME Data Marketplace from the FAME technical partners

Table 2 illustrates data that are generated by FAME during the project's lifetime.

Table 2 – Data collected and/or generated in FAME

Data Name	FAME Partner(s)
FAME Internal Data	
Data collected/generated through internal administration procedures	
Internal communication data	ALL
Partner's data	ALL
Internal administrative data	ALL
Data collected/generated for communication procedures	
Deliverables	ALL
Publications	ALL
Users' registration data	ALL
Data collected/generated for technical purposes	
Solution prototypes	FAME Technical Partners (WP2-WP5)
Software code	FAME Technical Partners (WP2-WP5)
Co-design activities data	FAME Technical & Pilot Partners (WP2-WP6)
Data collected/generated for FAME evaluation	
Evaluation data	FAME Technical & Pilot Partners (WP2-WP6)
Recommendations' data	FAME Technical & Pilot Partners (WP2-WP6)
FAME External Data	
Data (currently) indexed within the FAME Data Marketplace from the FAME pilots	

Pilot #2 – "Embedding Finance Services in a Personalized Citizen Wallet" Data assets	NOVO, DAEM
Pilot #5 – "ESG Scorecard Ranking & Sustainable Portfolio Optimisation" Data assets	KM, JRC
Pilot #6 – "Embedding Climatic Predictions in Property Insurance Products" Data assets	NRS, GFT
Pilot #7 – "Assessing the Quality and Monetary Value of Data Assets" Data assets	MOH, INNOV
Data (currently) indexed within the FAME Data Marketplace from the FAME technical partners	
Technical Data assets	FAME Technical Partners (WP2-WP5)

3.2 Formats of the Data

The collected and generated data are produced in various formats, depending on the nature and purpose of the data collected. Wherever possible, FAME intends to give preference to open, non-proprietary standards (including de facto standards), hence enhancing the data's interoperability and potential for reuse. The forms and formats in which the various FAME data can be found are shown in the following table (Table 3).

Table 3 – Data formats of the FAME data

Data Name	Data Format(s)
FAME Internal Data	
Data collected/generated through internal administration procedures	
Internal communication data	.msg, .pptx, .docx, .xlsx, .pdf
Partner's data	.docx, .xlsx
Internal administrative data	.docx, .pdf
Data collected/generated for communication procedures	
Deliverables	.docx, .pdf
Publications	.docx, .pdf, .latex
Users' registration data	Drupal database format
Data collected/generated for technical purposes	
Solution prototypes	.pptx, Adobe XD file format
Software code	.ts, .tsx, .mts, .cts, .java, .py, .ipynb, .c++
Co-design activities data	.jpeg, .pdf, formats derived from questionnaires, interviews, mock-ups, trade-off experiments
Data collected/generated for FAME evaluation	
Evaluation data	.docx, .pdf, .mp3
Recommendations' data	
FAME External Data	
Data (currently) indexed within the FAME Data Marketplace from FAME pilots	
Pilot #2 – "Embedding Finance Services in a Personalized Citizen Wallet" Data assets	.spss, .xlsx, .csv, .sav, .txt, .dat, .pdf, .docx, .xml, .glb, .png, .mp4, .json
Pilot #5 – "ESG Scorecard Ranking & Sustainable Portfolio Optimisation" Data assets	

Pilot #6 – "Embedding Climatic Predictions in Property Insurance Products" Data assets	
Pilot #7 – "Assessing the Quality and Monetary Value of Data Assets" Data assets	
Data (currently) indexed within the FAME Data Marketplace from FAME technical partners	
Technical Data assets	.ts, .tsx, .mts, .cts, .java, .py, .ipynb, .c++, .pptx, .docx, .xlsx, .pdf

3.3 Origin of the data

The origin of the data depends on each dataset. Some of these data were collected from partners and experts. Data was also created from scratch using different software tools or collected and combined directly through the data lake that was coming from the FAME pilots and the associated marketplaces. Based on the project's goals and to achieve synergies and maximize the impact as well as the total number of data being available throughout the FAME Data Marketplace, products/solutions from other EU projects, marketplaces and partners were leveraged and are listed on the table below (Table 4):

Table 4 – FAME Marketplaces and Data Spaces

Name	Description	Owner	Sector	Available Data
AGORA	AGORA is a data-marketplace connecting Data Providers and Data Consumers for selling and acquiring Connected Vehicle and Home Building data under the standardized data model (CIDM, Common Industrial Data Model). It offers a secure and privacy preserving experience when selling or buying sharing big data, by having the full control over your data shared, to whom and for what purposes. Even though AGORA is targeted mainly to vehicles and home building information, it can be also an implemented in several environments such as smart health or smart energy. AGORA offers to cross-sectorial Data Consumers to combine, integrate and analyse all data at once – regardless of source, type, size, or format – to generate the insights needed to address a wide range of end-users and customers' needs and transform them into new business opportunities in the era of Data Economy.	ATOS	Automotive (initially)	Open data sets
SecureIoT	The SecureIoT marketplace is a one stop-shop for Internet of Things (IoT) security solutions and an on-line community of IoT security solutions stakeholders. It facilitates access to IoT security solutions and knowledge, including datasets from the project testing phase.	Datasets owned by organization that uploaded them	Industry, IoT	Datasets from 3 use cases, webinars, presentations

IoT-Catalogue	Bringing IoT users and technology providers together from domain needs to IoT products (and back) via validated solutions with components, assembly guides and more.	UNP	IoT	115 Use Cases, 2119 Components
FINSEC	FINSEC (https://finsecurity.eu) training assets related to security aspects in the finance sector. The existing marketplace contains resources about Critical Infrastructure Protection	INNOV, GFT	Finance, Security	100 resources including Training Presentations, Webinars, Whitepapers, Workshops, presentations and training workshops
INFINITECH Marketplace	Solutions & Services for Finance. The INFINITECH marketplace was established to provide access to the project's solutions, accompanied a Virtualized Digital Innovation Hub (VDIH), aiming to support the innovators (FinTech/InsuranceTech) in their Big Data, AI, Blockchain and IoT endeavours.	INFINITECH Project	Finance	64 Total Assets (containing 18 Datasets, 16 ML algorithms and 8 blockchain assets among others), 85 3rd-Party Tools, 75 Accelerator Programmes, 2 Innov. Support Services, 213 Courses, 14 Workshops, 4 Hackathons, 4 Webinars, 6 How To Videos, 7 Whitepapers
INFINITECH-Zenodo	Data for the INFINITECH Project - Tailored IoT & BigData Sandboxes and Testbeds for Smart, Autonomous and Personalized Services in the European Finance and Insurance Services Ecosystem	INFINITECH Consortium	Finance	N/A
i3-MARKET	i3-MARKET develops a platform for interconnecting and federating heterogeneous marketplaces. It will therefore serve as a base component of the FAME platform: It will be enhanced with the project's AAI, analytics and data trading/monetization results.	i3-MARKET Project	Manufacturing, Climate, Automotive	50.000 datasets
Kaggle	Kaggle offers a no-setup, customizable, Jupyter Notebooks environment. Access GPUs at no cost to you and a huge repository of community published data & code.	Kaggle Inc., 1600 Amphitheatre Parkway, Mountain View, California 94043 USA	Finance, economy, industry, news, etc.	A few hundred datasets; an additional thousands of datasets via the connected Mobilithek, i.e., the national mobility data access point
Mobility Data Space	Mobility Data Space (Germany)	DRM Datenraum Mobilität GmbH	Mobility	87 data assets

PolicyCLOUD Data Marketplace	The Data Marketplace of the Policy Cloud project gives access to the Policy Cloud Data, Assets and Policies, for improving the modelling, creation and implementation of public and business decision making and policy, for developing common strategies and policies and, ultimately, for enhancing the citizens' quality of life.	UPRC	Smart Cities, Radicalization, Winery, Global Terrorism, Environment	
ADVANEO Data Marketplace	Advaneo has a vision to provide the most open possible access to data for all users serves as the basis for innovation and progress. Open data is a core aspect of the offering (around 2 million open data records from almost 20000 providers). The data marketplace serves as a data catalogue for the metadata provided by the data provider. In addition, it functions as a secure trading portal and offers functions for data processing and administration. It's a decentralized solution that gives users full control over their data at any time. This means that all relevant and possibly sensitive raw data always remain with the data provider and only is transferred directly – peer to peer – to the buyer in the event of a purchase. The data marketplace only stores the data descriptions (metadata) and has no access to your raw data at any time.	ADVANEO GmbH	Cross-sector	2 million Open datasets + potentially some commercial datasets (number not disclosed)

3.4 Size and Volume of the Data

The size of each data can be found in Section 3.6, for each different data type.

3.5 Data Utility

All the project's stakeholders may find one or more of the data collected or produced during FAME to be of particular value and may benefit from having access to and reusing this data. The tables created for each data that is anticipated to be collected or generated in the context of the project in the following Section (Section 3.6) provide more specific information on the stakeholder groups who may be interested in the reuse of each data.

3.6 FAME Data

3.6.1 FAME Internal Data

3.6.1.1 Data collected/generated through internal administration procedures

Table 5 – Internal communication data

Internal communication data	
Data description	This data is related with documentation of communication among members of the consortium (e.g., e-mails exchanged internally for FAME communication purposes, Meeting/Web conference related material)
Related WP/Task	ALL

Data origin	Communication among consortium partners
Will you re-use any existing data? If yes, how?	Materials produced in other parts of the project will be reused in the construction of this data
Methodologies for data collection/generation	Collaboration among project partners
Data format	The data will be provided in: .msg, .pptx, .docx, .xlsx, .pdf
Where this data will be stored?	Data will be stored internally by the project team in the password-protected document repository. Key outputs of these contributions are shared in the context of public deliverables or as events' reports
Expected size of the data	Order of Gigabytes
Metadata and Standards used	Each data source is accompanied by the following minimum set of metadata: (i) Title, (ii) Author, (iii) Creation date and purpose, and (iv) Usage - relevant group and meeting date
For whom might the data be useful?	This data will be useful for FAME partners
Data access, sharing and Licensing	The mailing list conversations and contacts of participants will remain confidential as required under GDPR

Table 6 – Partners' data

Partners' data	
Data description	This data is related with information and knowledge that will be gathered from questionnaires that will be shared to the consortium partners
Related WP/Task	WP2-WP6
Data origin	Replies to specific questionnaires
Will you re-use any existing data? If yes, how?	N/A
Methodologies for data collection/generation	Organization of webinars, and online training that will lead the process of collecting data through questionnaires
Data format	The data will be provided in: .docx, .xlsx
Where this data will be stored?	Data will be stored internally by the project team in the password-protected document repository. Key outputs of these contributions are shared in the context of public deliverables or as events' reports
Expected size of the data	Order of Megabytes
Metadata and Standards used	Sometime there was used the Schema.org organizational metadata standards
For whom might the data be useful?	This data will be useful for FAME partners
Data access, sharing and Licensing	N/A

Table 7 – Internal administrative data

Internal administrative data	
Data description	This data is related with administrative data that is collected internally among the consortium partners for record keeping and possibly statistical analysis
Related WP/Task	ALL
Data origin	Administrative information collected from partners
Will you re-use any existing data? If yes, how?	N/A
Methodologies for data collection/generation	Collaboration among project partners
Data format	The data will be provided in: .docx, .xlsx, .pdf
Where this data will be stored?	Data will be stored internally by the project team in the password-protected document repository
Expected size of the data	Order of Megabytes
Metadata and Standards used	Sometime there was used the Schema.org organizational metadata standards
For whom might the data be useful?	This data will be useful for FAME partners
Data access, sharing and Licensing	Internal, restrictive access available on case-by-case basis upon request

3.6.1.2 Data collected/generated for communication procedures

Table 8 – Deliverables

Deliverables	
Data description	This data is related with a report/document created as a result of FAME and intended for delivery to the EC
Related WP/Task	ALL
Data origin	Collaboration among FAME partners
Will you re-use any existing data? If yes, how?	Data originated from previous/on-going project activities of the involved partners in the form of baseline knowledge
Methodologies for data collection/generation	Collaboration among project partners in which each responsible for the deliverable (Task leader) will assign dedicated tasks to the involved partners
Data format	The data will be provided in: .docx, .pdf
Where this data will be stored?	Data will be stored internally by the project team in the password-protected document repository. It will be also shared in the https://www.fame-horizon.eu/ and Zenodo platform
Expected size of the data	Order of Gigabytes
Metadata and Standards used	Each data source is accompanied by the following minimum set of metadata: (i) Title, (ii) Author, (iii) Creation date and purpose, and (iv) Usage - relevant group and meeting date

For whom might the data be useful?	This data will be useful for FAME external stakeholders and project partners
Data access, sharing and Licensing	This data will become freely / openly accessible based on an open license. In the case that a deliverable will be marked as confidential, it will be only accessible from the project partners

Table 9 – Publications

Publications	
Data description	This data is related with a report/document intended to further the progress of science, usually by sharing findings from research with readers
Related WP/Task	ALL
Data origin	Collaboration among FAME partners
Will you re-use any existing data? If yes, how?	Data originated from previous/on-going project activities of the involved partners in the form of baseline knowledge
Methodologies for data collection/generation	Collaboration among project partners in which the involved publication authors will provide their research expertise into compiling the final document
Data format	The data will be provided in: .docx, .pdf, .latex
Where this data will be stored?	Data will be stored internally by the project team in the password-protected document repository. It will be also indexed in publishers' websites (e.g., Elsevier, Springer, IEEE) whereas it will be also shared in the https://www.fame-horizon.eu/ and Zenodo platform. In some cases, if it is appropriate, it will be shared in Open Access repositories (e.g., Institutional repositories, MedRxiv, European Open Science Cloud)
Expected size of the data	Order of Megabytes
Metadata and Standards used	Each data source will be accompanied by the sets of metadata that are required from the different publishers
For whom might the data be useful?	This data will be useful for FAME external stakeholders (e.g., scientific community, policymakers, AI/ML experts) and project partners
Data access, sharing and Licensing	This data will be accessible as it will be required from the different publishers' policies

Table 10 – Users' registration data

Users' registration data	
Data description	This data is related with personal data that are utilized to identify a registered user to the FAME website and related newsletter
Related WP/Task	WP7
Data origin	Data collected from https://www.fame-horizon.eu/contact-us/ and https://www.fame-horizon.eu/ (newsletter subscription)
Will you re-use any existing data? If yes, how?	N/A
Methodologies for data collection/generation	Registration form compiled by the user

Data format	The data will be provided in data format supported by the project's website
Where this data will be stored?	Data will be stored in the database (MySQL) of the related website
Expected size of the data	Order of Megabytes (less than 1 Megabyte per registered user)
Metadata and Standards used	Sometime there was used the Schema.org organizational metadata standards
For whom might the data be useful?	This data will be useful for the FAME consortium for the scope listed in the privacy policy statement
Data access, sharing and Licensing	Data will be accessed by the registered user, by the Data controller and by the Data Processor identified in the Privacy Policy

3.6.1.3 Data collected/generated for technical purposes

Table 11 – Solution prototypes

Solution prototypes	
Data description	This data is related with an early sample, model, or release of a product or system built to test a concept or assumption or to act as a thing to be tested and learned from
Related WP/Task	WP2, WP3, WP4, WP5
Data origin	Collaboration among FAME partners
Will you re-use any existing data? If yes, how?	Data originated from previous/on-going project activities of the involved partners in the form of baseline knowledge
Methodologies for data collection/generation	Collaboration among project partners in which each technical partner will provide his technical knowhow and background expertise
Data format	The data will be provided in: .pptx, Adobe XD file format
Where this data will be stored?	Data will be stored internally by the project team in a password-protected code repository – this will be determined as the project will become more mature
Expected size of the data	Order of Gigabytes
Metadata and Standards used	General code quality standards
For whom might the data be useful?	This data will be useful for FAME external stakeholders (e.g., scientific community, policymakers, AI experts) and technical project partners
Data access, sharing and Licensing	This data will become freely/openly accessible based on an open license by the time it will be considered appropriate

Table 12 – Software code

Software code	
Data description	This data is related with any collection of text, with or without comments, written using a human-readable programming language, usually as plain text
Related WP/Task	WP2, WP3, WP4, WP5

Data origin	Collaboration among FAME partners
Will you re-use any existing data? If yes, how?	Data originated from previous/on-going project activities of the involved partners in the form of baseline knowledge
Methodologies for data collection/generation	Collaboration among project partners in which each technical partner will provide his technical knowhow and background expertise
Data format	The data will be provided in: .ts, .tsx, .mts, .cts, .java, .py, .ipynb, .c++
Where this data will be stored?	Data will be stored internally by the project team in a password-protected code repository – this will be determined as the project will become more mature
Expected size of the data	Order of Gigabytes
Metadata and Standards used	General code quality standards
For whom might the data be useful?	This data will be useful for FAME external stakeholders (e.g., scientific community, policymakers, AI experts) and technical project partners
Data access, sharing and Licensing	This data will become freely/openly accessible based on an open license by the time it will be considered appropriate

Table 13 – Co-design activities data

Co-design activities data	
Data description	This data is related with data collected during co-design and co-creation sessions with Pilot partners
Related WP/Task	WP2, WP6
Data origin	Co-design sessions among FAME related partners
Will you re-use any existing data? If yes, how?	N/A
Methodologies for data collection/generation	Co-design sessions according to protocol described in T6.1 (WP6)
Data format	The data will be provided in: .jpeg, .pdf, .xlsx
Where this data will be stored?	Physically generated data (e.g., drawings, text) will be transformed to digital form and stored in the password-protected project's repository
Expected size of the data	Order of Gigabytes
Metadata and Standards used	Sometimes the ISO 9241 human-centred design metadata
For whom might the data be useful?	Designers/Developers of the FAME solutions
Data access, sharing and Licensing	This data will be accessible only to dedicated partners from the consortium involved in the co-design process

3.6.1.4 Data collected/generated for FAME evaluation

Table 14 – Evaluation data

Evaluation data	
Data description	Data for the evaluation of the proof-of-concept study of the FAME solutions, including measures of feasibility (i.e., acceptability, demand, implementation, practicality, adaptation, integration, and expansion), user experience and usability
Related WP/Task	WP6
Data origin	The data will be collected during the co-creation meetings with the focus groups of the pilot sites
Will you re-use any existing data? If yes, how?	N/A
Methodologies for data collection/generation	Data will be collected via surveys, focus groups and interviews with different users of the ecosystem at the pilot sites
Data format	The data will be provided in: .docx, .pdf, .mp3
Where this data will be stored?	Digital or electronic data will be stored on password-protected computers and secure servers hosted by the responsible for the co-creation methodologies
Expected size of the data	10 Gigabytes
Metadata and Standards used	Sometimes the ISO 9241 human-centred design metadata
For whom might the data be useful?	This data will be useful for all the project partners
Data access, sharing and Licensing	This has not been decided yet, since for the proof-of-concept studies, data deriving from different countries might need to be shared

Table 15 – Recommendations' data

Recommendations' data	
Data description	Data for the development of recommendations for a wider implementation of the FAME tools for personalised prevention strategies
Related WP/Task	WP6, WP7
Data origin	The data will be collected from focus group interviews with stakeholders involved in the proof-of-concept studies in the pilot sites. It will also include data from a questionnaire to collect input from a broader stakeholder group, with stakeholders from relevant national and international organizations and networks
Will you re-use any existing data? If yes, how?	N/A
Methodologies for data collection/generation	It will be used both the question-answering approach (e.g., Delphi study approach), as well interviews
Data format	The data will be provided in: .docx, .pdf, .mp3
Where this data will be	Digital or electronic data will be stored on password-protected

stored?	computers and secure servers hosted by the responsible for the co-creation methodologies
Expected size of the data	10 Gigabytes
Metadata and Standards used	FAIR-compliant policy recommendation metadata
For whom might the data be useful ?	This data will be useful for all the project partners
Data access, sharing and Licensing	These datasets will be only accessible to the relevant researchers on the project

3.6.2 FAME External Data

3.6.2.1 Data (currently) indexed within the FAME Data Marketplace from FAME pilots

3.6.2.1.1 Pilot 2 – "Embedding Finance Services in a Personalized Citizen Wallet" Data assets

Table 16 – Park & Ride Service

Park & Ride Service		
Data Asset Overview	**Description**	A service enabling NOVO's App users to purchase bus tickets directly through the citizen wallet when they park in designated areas. Integrates parking and public transport to promote sustainable mobility.
	Source/Origin	Generated by NOVO's mobility platform and integrated municipal parking systems.
	Collection/Generation methodology	Service generates transactional events when a user initiates a park session and purchases a bus ticket.
	Type	Service asset (transactional service)
	Format	JSON transaction logs, API calls, server-side events
	Behavior	Dynamic; updates in real time as citizens utilize the service.
	Expected size	Small to medium (dependent on usage volume, approx. thousands of transactions/day).
Metadata	**Metadata & Standards used**	DCAT-based metadata + FAME Marketplace metadata schema
	Documentation methodology	Automatically logged API documentation + internal service documentation.
Storage, Backup & Security	**Storage methodology**	Stored on NOVO's secure cloud servers with role-based access.
	Backup methodology	Daily backups with rolling 30-day retention.
	Security methodology	Encrypted channels (HTTPS/TLS), access logs, API key protection, GDPR-compliant data minimization.

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Contains pseudonymised transaction IDs; no direct personal data stored.
	Intellectual property rights responsible	NOVOMATIX PC
	Intellectual property rights constraints	Service is proprietary; shared only through licensing.
Access & Sharing	**Receiver(s)**	Municipalities, FAME Marketplace users.
	Re-use methodology	Service can be integrated via API into other city mobility systems
	Any other existing data (asset) re-use	Parking transaction dataset; OASA telematics dataset.
	Access, sharing & licensing methodology	Commercial license; shared via FAME Marketplace.
Archiving & Providing Access	**Preservation planning**	Long-term preservation of service documentation and API versions
	Long-term access provision planning	Service remains accessible through NOVO's API gateway.
	Preservation duration	2 years after project end
Responsibility	**Responsible for data asset management**	Novomatix PC

Table 17 – Loyalty Program Service

Loyalty Program Service		
Data Asset Overview	**Description**	Rewards users for frequent bus usage by offering free parking hours. Generates loyalty points and reward eligibility based on user behavior
	Source/Origin	Generated through NOVO's mobility/payment platform.
	Collection/Generation methodology	Events recorded when users validate bus rides or parking sessions.
	Type	Service asset.
	Format	Event logs (JSON), user reward states, API calls.
	Behavior	Dynamic, updates continuously per user activity
	Expected size	Moderate (dependent on adoption rate).
Metadata	**Metadata & Standards used**	DCAT-based metadata, FAME asset indexing fields
	Documentation methodology	Automatically logged API documentation + internal service documentation.

Storage, Backup & Security	**Storage methodology**	Stored on NOVO's secure cloud servers with role-based access
	Backup methodology	Daily backups with rolling 30-day retention
	Security methodology	Encrypted channels (HTTPS/TLS), access logs, API key protection, GDPR-compliant data minimization.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Only pseudonymised user identifiers; GDPR-compliant.
	Intellectual property rights responsible	NOVOMATIX PC.
	Intellectual property rights constraints	Service is proprietary; shared only through licensing.
Access & Sharing	**Receiver(s)**	Municipalities, FAME Marketplace users.
	Re-use methodology	Service can be integrated via API into other city mobility systems
	Any other existing data (asset) re-use	Parking transaction dataset; OASA telematics dataset.
	Access, sharing & licensing methodology	Commercial license; shared via FAME Marketplace.
Archiving & Providing Access	**Preservation planning**	Long-term preservation of service documentation and API versions
	Long-term access provision planning	Service remains accessible through NOVO's API gateway.
	Preservation duration	2 years after project end
Respon sibility	**Responsible for data asset management**	NOVOMATIX PC

3.6.2.1.2 Pilot 5 – "ESG Scorecard Ranking & Sustainable Portfolio Optimisation" Data assets

Table 18 – Aggressive ESG-Selective Portfolio

Aggressive ESG-Selective Portfolio		
Data Asset Overview	**Description**	Aggressive ESG selection investment portfolio focused on high-growth assets screened for ESG performance.
	Source/Origin	Internal portfolio management system; ESG scores sourced from external ESG data providers.
	Collection/Generation methodology	Extracted from internal databases; ESG metrics retrieved via API feeds and integrated
	Type	Structured financial and ESG dataset.
	Format	JSON

	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary with field definitions and lineage.
Storage, Backup & Security	**Storage methodology**	Stored in secure internal cloud repository.
	Backup methodology	Daily automated backups with version control.
	Security methodology	Encryption in transit and at rest; MFA and access logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No personal data; confidential investment information.
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access.
	Re-use methodology	Used for ESG reporting, performance analytics, investment screening.
	Any other existing data (asset) re-use	May integrate with ESG scores, benchmarks, historical records.
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA; not licensed externally.
Archiving & Providing Access	**Preservation planning**	Archived annually.
	Long-term access provision planning	Maintained in long-term archival storage.
	Preservation duration	Minimum 10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 19 – Aggressive Growth Portfolio

Aggressive Growth Portfolio		
Data Asset Overview	**Description**	Aggressive growth investment portfolio, including asset allocation, performance metrics, and risk attributes.
	Source/Origin	Internal portfolio management system; supplemented by external market data providers.
	Collection/Generation methodology	Extracted from internal investment databases; market data collected via automated API feeds.
	Type	Structured financial dataset.
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented through internal data dictionary, describing fields, field types, update frequency, and data lineage.
Storage, Backup & Security	**Storage methodology**	Stored in a secure internal cloud repository with access control policies applied.
	Backup methodology	Daily automated backups to a secondary secure environment with versioning enabled.
	Security methodology	Access restricted via role-based access control, encryption at rest and in transit, multi-factor authentication
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Contains no personal data; confidentiality relates only to proprietary investment information.
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Data may be reused for performance reporting, risk analysis, ESG scoring integration, and trend comparisons.
	Any other existing data (asset) re-use	May incorporate or correlate with market benchmarks, ESG datasets, and historical

		portfolio records.
	Access, sharing & licensing methodology	Access granted through internal data request process; shared under internal data-use policy; not licensed externally.
Archiving & Providing Access	**Preservation planning**	Archived annually to maintain historical record of portfolio configurations.
	Long-term access provision planning	Accessible through the organization's long-term archival system with guaranteed compatibility for future retrieval.
	Preservation duration	Minimum of 10 years, aligned with regulatory and audit requirements.
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 20 – Aggressive Sustainable Portfolio

Aggressive Sustainable Portfolio		
Data Asset Overview	**Description**	Aggressive sustainable portfolio with focus on high-growth assets meeting sustainability criteria
	Source/Origin	Internal portfolio system; external sustainability data providers
	Collection/Generation methodology	Internal extracts plus sustainability KPIs pulled via APIs
	Type	structured
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud storage; daily backups; encryption
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential investment data; no personal data
	Intellectual property rights responsible	KM & JRC

	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Reused for sustainability performance reporting
	Any other existing data (asset) re-use	Integration with ESG scores, benchmarks, historical records
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA; not licensed externally
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 21 – Balanced ESG-Aware Portfolio

Balanced ESG-Aware Portfolio		
Data Asset Overview	**Description**	Balanced ESG-aware portfolio combining moderate-risk allocation with ESG integration
	Source/Origin	Internal system + external ESG indicators
	Collection/Generation methodology	Internal extracts with ESG metric integration
	Type	Financial + ESG structured dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary

Storage, Backup & Security	**Storage methodology**	Secure cloud, daily backups, encryption, RBAC
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential portfolio info
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	ESG scoring and reporting
	Any other existing data (asset) re-use	Integration with ESG data and benchmarks
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 22 – Balanced Impact Portfolio

Balanced Impact Portfolio		
Data Asset Overview	**Description**	Balanced impact portfolio targeting measurable positive environmental and social impact
	Source/Origin	Internal portfolio data; external impact KPIs
	Collection/Generation methodology	Internal investment data integrated with impact metrics from verified providers
	Type	Financial + impact measurement dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine

	Expected size	≈22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Cloud storage; daily backups; encryption & RBAC
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No personal data; proprietary investment and impact data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Impact and investment reporting
	Any other existing data (asset) re-use	Integration with ESG or sustainability data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 23 – Balanced Standard Portfolio

Balanced Standard Portfolio		
Data Asset Overview	**Description**	Balanced standard portfolio without specific ESG or sustainability adjustments
	Source/Origin	Internal portfolio system; market data providers
	Collection/Generation methodology	Financial data extracted from internal systems; market feeds integrated automatically
	Type	Structured financial dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure internal cloud; daily backups
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential proprietary portfolio data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Investment reporting and analysis
	Any other existing data (asset) re-use	Integration with market data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providin	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage

	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 24 – Conservative Conscious Portfolio

Conservative Conscious Portfolio		
Data Asset Overview	**Description**	Conservative conscious portfolio incorporating low-risk allocation with basic ESG considerations
	Source/Origin	Internal system; basic ESG provider inputs
	Collection/Generation methodology	Internal extracts combined with light ESG scoring
	Type	Structured financial + minimal ESG dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Cloud with RBAC; daily backups
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential investment dataset
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Conservative ESG reporting
	Any other existing data (asset) re-use	Integration with minimal ESG scoring data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial

		licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 25 – Conservative Sustainable Portfolio

Conservative Sustainable Portfolio		
Data Asset Overview	**Description**	Conservative sustainable portfolio focusing on low-risk, sustainability-oriented investments
	Source/Origin	Internal portfolio records; sustainability data vendors
	Collection/Generation methodology	Extracted internally; sustainability KPIs integrated as needed
	Type	Financial + sustainability structured dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud environment; daily backup routine
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential investment data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse

Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Sustainability reporting
	Any other existing data (asset) re-use	Integration with ESG and sustainability KPIs
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual snapshots
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 26 – Conservative Traditional Portfolio

Conservative Traditional Portfolio		
Data Asset Overview	**Description**	Conservative traditional portfolio focusing on low-risk assets without ESG or sustainability criteria
	Source/Origin	Internal investment system; market data providers
	Collection/Generation methodology	Financial records extracted internally; market feeds integrated
	Type	Structured financial dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈22KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud storage, backups, encryption
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential proprietary data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Investment and risk reporting
	Any other existing data (asset) re-use	Integration with market data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 27 – Ranking Aggressive ESG-Selective Portfolio

Ranking Aggressive ESG-Selective Portfolio		
Data Asset Overview	**Description**	ESG ranking of high-growth aggressive portfolio assets
	Source/Origin	Internal portfolio management system; ESG data providers
	Collection/Generation methodology	Computed internally using ESG scores and performance metrics
	Type	Structured financial and ESG ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine snapshot for reporting
	Expected size	≈11KB
Meta data	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory

	Documentation methodology	Documented via internal data dictionary with field definitions and lineage
Storage, Backup & Security	**Storage methodology**	Secure internal cloud repository with RBAC
	Backup methodology	Daily automated backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No personal data; confidential ranking information
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Investment screening, performance analytics, ESG reporting
	Any other existing data (asset) re-use	May integrate with portfolio historical records, benchmarks
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA; not licensed externally
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 28 – Ranking Aggressive Growth Portfolio

Ranking Aggressive Growth Portfolio		
Data Asset Overview	**Description**	ESG ranking of high-growth aggressive portfolio assets
	Source/Origin	Internal portfolio management system; ESG data providers
	Collection/Generation methodology	Computed internally using ESG scores and performance metrics
	Type	Structured financial and ESG ranking dataset
	Format	JSON

	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine snapshot for reporting
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary with field definitions and lineage
Storage, Backup & Security	**Storage methodology**	Secure internal cloud repository with RBAC
	Backup methodology	Daily automated backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No personal data; confidential ranking information
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Investment screening, performance analytics, ESG reporting
	Any other existing data (asset) re-use	May integrate with portfolio historical records, benchmarks
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA; not licensed externally
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 29 – Ranking Aggressive Sustainable Portfolio

Ranking Aggressive Sustainable Portfolio		
Data Asset Overview	**Description**	ESG and sustainability ranking of high-growth aggressive sustainable portfolio assets
	Source/Origin	Internal portfolio system; sustainability data providers
	Collection/Generation methodology	Computed internally using sustainability KPIs and ESG scores
	Type	Structured financial & sustainability ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine snapshot for reporting
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud storage; daily backups; encryption and RBAC
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential ranking information; no personal data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	ESG and sustainability reporting
	Any other existing data (asset) re-use	Integration with historical rankings and benchmarks
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA; not licensed externally

Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 30 – Ranking Balanced ESG-Aware Portfolio

Ranking Balanced ESG-Aware Portfolio		
Data Asset Overview	**Description**	ESG ranking of balanced ESG-aware portfolio assets
	Source/Origin	Internal system + external ESG indicators
	Collection/Generation methodology	Computed internally using ESG integration metrics
	Type	Financial + ESG ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud, daily backups, encryption, RBAC
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential ranking information
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	ESG scoring and reporting

	Any other existing data (asset) re-use	Integration with ESG data and benchmarks
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 31 – Ranking Balanced Impact Portfolio

Ranking Balanced Impact Portfolio		
Data Asset Overview	**Description**	Impact ranking of balanced impact portfolio assets targeting measurable environmental and social outcomes
	Source/Origin	Internal portfolio data; external impact KPIs
	Collection/Generation methodology	Computed internally using impact KPIs and investment data
	Type	Financial + impact measurement ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Cloud storage; daily backups; encryption & RBAC
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No personal data; proprietary ranking information
	Intellectual property rights responsible	KM & JRC

	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Impact and investment reporting
	Any other existing data (asset) re-use	Integration with ESG or sustainability data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 32 – Ranking Balanced Standard Portfolio

Ranking Balanced Standard Portfolio		
Data Asset Overview	**Description**	Financial ranking of balanced standard portfolio assets without ESG or sustainability adjustments
	Source/Origin	Internal portfolio system; market data providers
	Collection/Generation methodology	Computed internally using financial metrics and market data
	Type	Structured financial ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary

Storage, Backup & Security	**Storage methodology**	Secure internal cloud; daily backups
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential proprietary ranking data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Investment and risk reporting
	Any other existing data (asset) re-use	Integration with market data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 33 – Ranking Conservative Conscious Portfolio

Ranking Conservative Conscious Portfolio		
Data Asset Overview	**Description**	ESG ranking of conservative conscious portfolio assets incorporating low-risk allocation
	Source/Origin	Internal system; basic ESG provider inputs
	Collection/Generation methodology	Computed internally using light ESG scoring and portfolio data
	Type	Structured financial + minimal ESG ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine

	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Cloud with RBAC; daily backups
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential ranking dataset
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Conservative ESG ranking reporting
	Any other existing data (asset) re-use	Integration with minimal ESG scoring data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 34 – Ranking Conservative Sustainable Portfolio

Ranking Conservative Sustainable Portfolio		
Data Asset Overview	**Description**	Sustainability ranking of conservative sustainable portfolio assets focusing on low-risk investments
	Source/Origin	Internal portfolio records; sustainability data vendors

	Collection/Generation methodology	Computed internally using sustainability KPIs and portfolio data
	Type	Financial + sustainability ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud environment; daily backup routine
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential ranking dataset
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Sustainability ranking reporting
	Any other existing data (asset) re-use	Integration with ESG and sustainability KPIs
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing Access	**Preservation planning**	Annual snapshots
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 35 – Ranking Conservative Traditional Portfolio

Ranking Conservative Traditional Portfolio		
Data Asset Overview	**Description**	Ranking of conservative traditional portfolio assets focusing on low-risk investments without ESG or sustainability criteria
	Source/Origin	Internal investment system; market data providers
	Collection/Generation methodology	Computed internally using financial metrics and market data
	Type	Structured financial ranking dataset
	Format	JSON
	Behavior	Fixed snapshot; regenerated monthly via Pilot 5 optimization engine
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud storage, backups, encryption
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Confidential proprietary ranking data
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	Internal KM/JRC teams and FAME Marketplace participants (e.g., Asset Managers, FinTechs) via subscription/token access
	Re-use methodology	Investment and risk reporting
	Any other existing data (asset) re-use	Integration with market data
	Access, sharing & licensing methodology	Access granted via FAME Marketplace (Smart Contract/Token authorization). Commercial licensing (Pay-per-use/Subscription) as per DoA
Archiving & Providing	**Preservation planning**	Annual archive
	Long-term access provision planning	Maintained in long-term archival storage

	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

Table 36 – Service Asset: Pilot 5 ESG Ranking & Portfolio Service (PaaS)

Service Asset: Pilot 5 ESG Ranking & Portfolio Service (PaaS)		
Data Asset Overview	**Description**	A Platform-as-a-Service (PaaS) model with a generative AI-enhanced optimization engine for real-time ESG scoring and portfolio optimization
	Source/Origin	KM/JRC Internal Development
	Collection/Generation methodology	Computed internally using financial metrics and market data
	Type	Service / API
	Format	REST API (JSON responses)
	Behavior	Dynamic / On-demand
	Expected size	≈11KB
Metadata	**Metadata & Standards used**	S&P 100 Index, AI Financial Metrics engine, ESG Scoring Framework, Modern Portfolio Theory.
	Documentation methodology	Documented via internal data dictionary
Storage, Backup & Security	**Storage methodology**	Secure cloud storage, backups, encryption
	Backup methodology	Daily backups with version control
	Security methodology	Encryption in transit and at rest; MFA and access logging
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Proprietary KM/JRC asset; available via FAME Marketplace.
	Intellectual property rights responsible	KM & JRC
	Intellectual property rights constraints	Proprietary KM/JRC asset made available via the FAME Marketplace for regulated reuse
Access & Sharing	**Receiver(s)**	FAME Marketplace users (Asset Managers, FinTechs).
	Re-use methodology	Investment and risk reporting
	Any other existing data (asset) re-use	Integration with market data
	Access, sharing & licensing methodology	Token-based access via FAME AAI; Subscription model.

Archiving & Providing Access	**Preservation planning**	Service logic maintained in code repositories
	Long-term access provision planning	Maintained in long-term archival storage
	Preservation duration	10 years
Respon sibility	**Responsible for data asset management**	KM & JRC

3.6.2.1.3 Pilot 6 – "Embedding Climatic Predictions in Property Insurance Products" Data assets

Table 37 – Extreme heat projections, CCCma-CanESM2/CCCma-CanRCM4

Extreme heat projections, CCCma-CanESM2/CCCma-CanRCM4		
Data Asset Overview	**Description**	Number of days per year with daily mean temperature above 30°C, calculated from CCCma-CanESM2/CCCma-CanRCM4 climate model simulations at 20,640 different locations in California.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-02-06)
	Collection/Generation methodology	See report linked below for details on how the extreme heat index projections were derived from the climate model output.
	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	8,8 MB
Metadat a	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212709
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup

		procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.
Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.
Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 38 – Extreme heat projections, CCCma-CanESM2/ OURANOS-CRCM5

Extreme heat projections, CCCma-CanESM2/ OURANOS-CRCM5		
Data Asset Overview	**Description**	Number of days per year with daily mean temperature above 30°C, calculated from CCCma-CanESM2/OURANOS-CRCM5 climate model simulations at 20,640 different locations in California.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-02-10)
	Collection/Generation methodology	See report linked below for details on how the extreme heat index projections were derived from the climate model output.
	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	8,9 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212709
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.
Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.
Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 39 – Extreme heat projections, CNRM-CERFACS-CM5/ OURANOS-CRCM5

Extreme heat projections, CNRM-CERFACS-CM5/ OURANOS-CRCM5		
Data Asset Overview	**Description**	Number of days per year with daily mean temperature above 30°C, calculated from CNRM-CERFACS-CM5/OURANOS-CRCM5 climate model simulations at 20,640 different locations in California.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from:

		Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-02-10)
	Collection/Generation methodology	See report linked below for details on how the extreme heat index projections were derived from the climate model output.
	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	8,6 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212709
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.

Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.
Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 40 – Extreme heat projections, MPI-M-MPI-ESM-LR/ OURANOS-CRCM5

Extreme heat projections, MPI-M-MPI-ESM-LR/ OURANOS-CRCM5		
Data Asset Overview	**Description**	Number of days per year with daily mean temperature above 30°C, calculated from MPI-M-MPI-ESM-LR/OURANOS-CRCM5 climate model simulations at 20,640 different locations in California.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-02-10)
	Collection/Generation methodology	See report linked below for details on how the extreme heat index projections were derived from the climate model output.

	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	8,5 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212709
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.
Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.

Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 41 – Extreme heat projections, NOAA-GFDL-ESM2M/OURANOS-CRCM5

Extreme heat projections, NOAA-GFDL-ESM2M/OURANOS-CRCM5		
Data Asset Overview	**Description**	Number of days per year with daily mean temperature above 30°C, calculated from NOAA-GFDL-ESM2M/OURANOS-CRCM5 climate model simulations at 20,640 different locations in California.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-02-10)
	Collection/Generation methodology	See report linked below for details on how the extreme heat index projections were derived from the climate model output.
	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	8,6 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212709

Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.
Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.
Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).

Responsibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 42 – Storm loss projections, CNRM-CERFACS-CM5/CNRM-ALADIN63

Storm loss projections, CNRM-CERFACS-CM5/CNRM-ALADIN63		
Data Asset Overview	**Description**	Projections of a storm loss index over Europe (divided into seven sub-regions), calculated from CNRM-CERFACS-CM5/CNRM-ALADIN63 climate model simulations.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-06-25)
	Collection/Generation methodology	See report linked below for details on how the storm loss index projections were derived from the climate model output.
	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	2,5 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212710
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status,

		providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.
Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.
Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 43 – Storm loss projections, ICHEC-EC-EARTH/DMI-HIRHAM5

Storm loss projections, ICHEC-EC-EARTH/DMI-HIRHAM5		
Data Asset Overview	**Description**	Projections of a storm loss index over Europe (divided into seven sub-regions), calculated from ICHEC-EC-EARTH/DMI-HIRHAM5 climate model simulations.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded

		from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-06-25)
	Collection/Generation methodology	See report linked below for details on how the storm loss index projections were derived from the climate model output.
	Type	Value-added climate model output
	Format	csv
	Behavior	fixed
	Expected size	2,4 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212710
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.

Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.
Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

Table 44 – Storm loss projections, MPI-M-MPI-ESM-LR/SMHI-RCA4

Storm loss projections, MPI-M-MPI-ESM-LR/SMHI-RCA4		
Data Asset Overview	**Description**	Projections of a storm loss index over Europe (divided into seven sub-regions), calculated from MPI-M-MPI-ESM-LR/SMHI-RCA4 climate model simulations.
	Source/Origin	The climate model simulations underlying our extreme heat index series were downloaded from: Copernicus Climate Change Service, Climate Data Store, (2019): CORDEX regional climate model data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.bc91edc3 (Accessed on 2025-06-25)
	Collection/Generation methodology	See report linked below for details on how the storm loss index projections were derived from the climate model output.
	Type	Value-added climate model output

	Format	csv
	Behavior	fixed
	Expected size	2,1 MB
Metadata	**Metadata & Standards used**	Does not apply
	Documentation methodology	See report at https://nr.brage.unit.no/nr-xmlui/handle/11250/3212710
Storage, Backup & Security	**Storage methodology**	Data storage in dedicated filesystem directory with unique UUID-based filenames; asset metadata managed in a separate database; SHA-256 hash for integrity verification
	Backup methodology	Data are hosted on cloud infrastructure; backup policies follow standard on-demand backup procedures provided by AWS
	Security methodology	Organization-based authentication with password hashing; session management with configurable timeouts; HTTPS enforced in production; secure session cookies configured appropriately. Download attempts comprehensively logged with timestamp, IP address, user agent, and success/failure status, providing a complete audit trail; Downloads go through the application layer with logging.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	No privacy or confidentiality issues
	Intellectual property rights responsible	The property rights remain with the creators of the underlying climate model simulations. See CORDEX license statement at 10.24381/cds.bc91edc3 for details.
	Intellectual property rights constraints	The license agreement linked above limits the use of the climate model simulations and data assets derived from them (like this one) to non-commercial research and educational purposes.
Access & Sharing	**Receiver(s)**	Real estate and insurance sector, research
	Re-use methodology	Access to this asset is provided through the FAME marketplace. The report linked above (Documentation methodology) describes a possible use case for this data asset (projecting the decrease of house values due to climate change.
	Any other existing data (asset) re-use	-
	Access, sharing & licensing methodology	Every organization registered with the FAME marketplace has access to this data asset. Access is free, but the intellectual property rights constraints mentioned above apply.

Archiving & Providing Access	**Preservation planning**	The data asset is fixed and will exist on the FAME marketplace for as long as the marketplace itself is maintained.
	Long-term access provision planning	Long-term access that goes beyond the existence of the FAME marketplace cannot be guaranteed.
	Preservation duration	A local copy of the data set will be maintained by the responsible for data asset management (see below) for three years (i.e. until 31.12.2028).
Respon sibility	**Responsible for data asset management**	Alex Lenkoski Norwegian Computing Center lenkoski@nr.no

3.6.2.1.4 Pilot 7 – "Assessing the Quality and Monetary Value of Data Assets" Data assets

Table 45 – MOH Compressor Sensor Datasets

MOH Compressor Sensor Datasets		
Data Asset Overview	**Description**	Raw time-series sensor data from MOH compressors (temperature, pressure, vibration, position). Enables analytics, predictive maintenance, and anomaly detection.
	Source/Origin	Collected by MOH refinery sensor network DCS system.
	Collection/Generation methodology	Automatic IIoT data capture; periodic sampling and aggregation via historian database.
	Type	Time-series dataset.
	Format	CSV, JSON, proprietary historian DB format.
	Behavior	Changing over time – updated continuously.
	Expected size	Approx. 500 MB/month (variable by sampling).
Metadata	**Metadata & Standards used**	Sensor ID, timestamp, calibration, machine, units; ISA/IEC 62443 principles.
	Documentation methodology	Automated documentation, machine/sensor labels, calibration reports.
Storage, Backup & Security	**Storage methodology**	Stored in MOH historian, mirrored to local secure servers, access via API.
	Backup methodology	Daily backup (historian), weekly mirrored copy to secure offsite.
	Security methodology	Encrypted transmission; access control; logging; audit trails.

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Privacy & Confidentiality issues | Industrial/operational confidentiality; not personally identifying.
	Intellectual property rights responsible	Motor Oil Hellas
	Intellectual property rights constraints	Cannot publish raw data externally without FAME/MOH approval.
Access & Sharing	**Receiver(s)**	Project partners (FAME), external users with permission.
	Re-use methodology	Accessible via FAME marketplace APIs/portal; suitable licensing.
	Any other existing data (asset) re-use	Reuses earlier refinery sensor baselines, cross-validation.
	Access, sharing & licensing methodology	Access under negotiated FAME licenses; usage restrictions apply.
Archiving & Providing Access	**Preservation planning**	Retained on secured servers beyond project duration.
	Long-term access provision planning	Documented procedures, retained as exemplar for future analytics.
	Preservation duration	Minimum 10 years.
Respon sibility	**Responsible for data asset management**	Aristotelis Ntafalias; Motor Oil Hellas; antafalias@moh.gr

3.6.2.2 Data (currently) indexed within the FAME Data Marketplace from FAME technical partners

Table 46 – Pricing Advisory Tool

Pricing Advisory Tool		
Data Asset Overview	**Description**	The PAT (Pricing Advisory Tool) product provides price recommendations based on user responses from a questionnaire, the selected business model, and the number of completed transactions of similar assets. Similar assets are identified based on the similarity of questionnaire answers and the similarity of individual asset descriptions. The database stores offerings created through the offering page (from the PT module), user answers from the questionnaire, and transaction information obtained from similar assets through recommendations from the TM module. The offerings also contain asset-related information such as description, title, and type, which are retrieved from FDAC (catalogue).

	Source/Origin	The data originate from user questionnaire inputs, internal PT and TM system modules, and external asset metadata sources provided by FDAC (catalogue).
	Collection/Generation methodology	The data are generated through a combination of manual user input via questionnaires and automated data processing of transaction records based on similar assets. Offerings are created using the internal PT module.
	Type	The data asset consists of database records including transaction data, questionnaire answers, offerings, and asset metadata stored as structured documents.
	Format	All data are stored in JSON format within a MongoDB database.
	Behavior	The data are dynamic and continuously updated with each new offering, user submission, and change in asset-related information.
	Expected size	The expected data size depends on the number of users, offerings, and stored transactions. The estimated storage volume is expected to range from several gigabytes to tens of gigabytes.
Metadata	**Metadata & Standards used**	The data asset uses a custom internal metadata schema defined within the MongoDB structure. Metadata include structured field names, data types, timestamps, asset identifiers, offering identifiers, and transaction references. Standard JSON data formatting is applied. No external international metadata standards are currently enforced.
	Documentation methodology	The data asset is documented through internal technical documentation, including system architecture descriptions, database structure documentation, and process descriptions in textual form. The documentation is maintained manually by the development team and updated when significant changes occur in the system or data structure.
Storage, Backup & Security	**Storage methodology**	The data are stored in a MongoDB database. Access to the data is restricted to authorized internal users only.
	Backup methodology	Data backup is planned to be fully automated and stored in a secure cloud environment.
	Security methodology	The data are intended strictly for internal use and are not shared externally. Access control mechanisms are applied to prevent unauthorized access.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	All collected data are handled in compliance with internal data protection policies. User-provided questionnaire data and transaction-related information are treated as confidential. Access to data is strictly limited to authorized internal users. No personal data are

		shared with external parties.
	Intellectual property rights responsible	The intellectual property rights for the PAT system and its data assets are owned by the developing organization. Ownership and usage rights are defined by internal company agreements.
	Intellectual property rights constraints	The data asset is restricted to internal use only. Any external use, sharing, publishing, or redistribution of the data is prohibited unless explicitly authorized by the data owner.
Access & Sharing	**Receiver(s)**	The primary end-users of the data are internal teams involved in pricing strategy, analytics, development, and business operations. No external research community access is currently planned.
	Re-use methodology	The data may be reused internally for pricing optimization, analytics, reporting, and model validation. External access is not permitted. Internal users access the data through controlled system modules and internal tools.
	Any other existing data (asset) re-use	The system reuses existing asset metadata retrieved from FDAC (catalogue) for enriching offerings with asset descriptions, titles, and types.
	Access, sharing & licensing methodology	Access to the data is granted based on internal authorization roles. No public platforms are used for data sharing. The data are not shared under any open license and remain proprietary.
Archiving & Providing Access	**Preservation planning**	Data will continue to be preserved within the production database and cloud backups after the project's active development phase. Long-term preservation will rely on regular backups and controlled access storage.
	Long-term access provision planning	Future internal users will be able to access the data through the existing system infrastructure and internal documentation that describes database structure and access procedures.
	Preservation duration	Data are planned to be retained for a long-term period, with a minimum retention duration of several years, depending on internal company policies and legal requirements.

Responsibility	**Responsible for data asset management**	The responsibility for data asset management lies with the internal development and data management team of the organization. **Responsible entity:** Name Surname: Jana Péliová Organization:EUBA Email: jana.peliova@euba.sk

Table 47 – IIoT Data Quality Assessment Service

IIoT Data Quality Assessment Service		
Data Asset Overview	**Description**	A complete web-based software system for assessing industrial IoT sensor data quality. Comprises FastAPI backend with RESTful API and AI agent, React/TypeScript frontend with interactive dashboards, LeanXcale database schema with online aggregates, Docker deployment infrastructure, and documentation. Enables plant operators to detect missing values, identify threshold violations, and assess data completeness through natural language queries and visualizations.
	Source/Origin	Developed by INNOV and validated by Pilot 7 (MOH) leveraging FAME's energy efficient analytics (LXS). Repository: https://github.com/giorgosfatouros/IIoT-Data-Quality-Assessment. Built using FastAPI, React, LeanXcale database integration, and OpenAI Agents SDK.
	Collection/Generation methodology	Agile software development with Git version control. Requirements gathered from MOH's engineers operating compressor stations. Code developed manually with AI assistance, tested and validated through real sensor datasets provided by MOH. Documentation authored alongside code with API specifications.
	Type	Software Application (Full-Stack Web Service).
	Format	Source: Python (.py), TypeScript (.tsx/.ts), JSON, YAML, Markdown (.md), Shell scripts (.sh).; Docker Image
	Behavior	*Changing*
	Expected size	Source code: ~6 MB. Dependencies: ~700 MB (Python + Node.js packages).

Metadata	**Metadata & Standards used**	OpenAPI 3.0, Dublin Core for documentation, Software Heritage, GIT Version control
	Documentation methodology	C code-level (docstrings, type hints, inline comments); API-level (Swagger/ReDoc); user-level (README files, setup guides); developer-level (deployment scripts, troubleshooting).
Storage, Backup & Security	**Storage methodology**	Git repository on GitHub (https://github.com/giorgosfatouros/IIoT-Data-Quality-Assessment)
	Backup methodology	Git ensures every clone is a backup
	Security methodology	No credentials in source code. Secret scanning via GitHub/Dependabot. Dependency vulnerability scanning with automated security updates. SQL injection prevented via SQLAlchemy ORM. Input validation through Pydantic models.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Software contains no personal data. User identifiers and raw data are not included in the AI agent context (the agent can be served by a local LLM). Industrial sensor data processed by deployed system remains confidential with data owners. GDPR minimal applicability (software itself, not operational data).
	Intellectual property rights responsible	Original code owned by INNOV, LeanXcale integration jointly with LXS, open-source dependencies retain original licenses (MIT, Apache 2.0, BSD).
	Intellectual property rights constraints	The code released as open source, with Apache 2.0 as the default licence. Any non-open components (LeanXcale) remain optional external dependencies. Commercial activity will focus on B2B deployment, integration and support rather than restricting software rights.
Access & Sharing	**Receiver(s)**	FAME consortium partners, open-source developers, industrial adopters (manufacturing facilities, IoT service providers, system integrators), technology evaluators (pilot participants) and educational institutions.
	Re-use methodology	Direct deployment via Docker following README. Customization by extending machine types, analytics functions, or UI components. API integration from external applications. Forking permitted under license terms.
	Any other existing data (asset) re-use	LeanXcale database system, MOH Importer Java application (companion FAME

		component), open-source frameworks specified in requirements.txt and package-lock.json.
	Access, sharing & licensing methodology	FAME Marketplace and Public GitHub repository with Apache 2.0 license for community edition, plus commercial license option for enterprise features and support.
Archiving & Providing Access	**Preservation planning**	Source code preserved indefinitely via GitHub. Documentation in repository (Markdown) and PDF archives.
	Long-term access provision planning	Software Heritage ID for commits, GitHub permanent URLs
	Preservation duration	Indefinitely: Source code, versioned releases, documentation, dependencies (lock files)
Respon sibility	**Responsible for data asset management**	George Fatouros; Innov-Acts; gfatouros@innov-acts.com

Table 48 – CO2 Emissions Monitoring API

CO2 Emissions Monitoring API		
Data Asset Overview	**Description**	A Flask API monitoring power consumption and calculating CO2 emissions from Kubernetes containers using Kepler metrics. Supports real-time and historical analysis.
	Source/Origin	GitHub Repository: E3-JSI/co2-emissions-prediction. Developed by the Department for Artificial Intelligence at the Jozef Stefan Institute.
	Collection/Generation methodology	Scrapes Prometheus metrics (Kepler) and CO2 intensity data
	Type	Software library, API
	Format	Python (.py), Dockerfile, Kubernetes YAML, SQL (TimescaleDB interaction).
	Behavior	Dynamic: Real-time monitoring application. Code is version controlled.
	Expected size	Docker Image (hundreds of MBs). Database storage depends on retention period.
Metadata	**Metadata & Standards used**	requirements.txt for Python env. Docker labels. Kubernetes deployment standards.
	Documentation methodology	Extensive README.md detailing API endpoints (POST, GET), environment variables, and deployment steps.
Storage, Backup & Security	**Storage methodology**	Code on GitHub. Data persistence via TimescaleDB (if configured) or in-memory mock.
	Backup methodology	Git for code.

	Security methodology	/
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Monitors container performance, non-personal data
	Intellectual property rights responsible	Jozef Stefan Institute (Dept. for Artificial Intelligence).
	Intellectual property rights constraints	Licensed under MIT License.
Access & Sharing	**Receiver(s)**	DevOps Engineers, Green Computing Researchers, Kubernetes Administrators.
	Re-use methodology	Deployable via Docker/Kubernetes. API can be integrated into monitoring dashboards
	Any other existing data (asset) re-use	/
	Access, sharing & licensing methodology	Public GitHub repository. MIT License.
Archiving & Providing Access	**Preservation planning**	Hosted on GitHub
	Long-term access provision planning	Source code availability via Git
	Preservation duration	Indefinite.
Responsibility	**Responsible for data asset management**	Department for Artificial Intelligence, Jozef Stefan Institute. Contributors: Ivo Hrib, Jan Šturm

Table 49 – StreamStory Causal (SSCausal)

StreamStory Causal (SSCausal)		
Data Asset Overview	**Description**	A Python framework that generates transaction data from a multivariate time series using the StreamStory Markov chain model. It includes hybrid feature selection, causal discovery capabilities, and LLM integration for labelling.
	Source/Origin	GitHub Repository: FAME-JSI/ss-causal. Developed by the Department for Artificial Intelligence at the Jozef Stefan Institute.
	Collection/Generation methodology	Generates transaction data for Causal Discovery from input time series. Uses variance, correlation, and SVD-based feature selection combined with StreamStory Markov models.
	Type	Software library
	Format	Python (.py), YAML (configuration), Shell scripts (.sh), CSV (input/output data), JSON (LLM descriptions).

	Behavior	Dynamic
	Expected size	Repository: Lightweight. Generated Data: Varies based on experiment configuration
Metadata	**Metadata & Standards used**	pyproject.toml for Python dependencies/versioning. YAML for experiment configuration standards.
	Documentation methodology	Documented via README.md, inline code comments, and CLI help commands.
Storage, Backup & Security	**Storage methodology**	Source code stored on GitHub. Local execution stores outputs in ./output/ directory (models, transactions).
	Backup methodology	Git version control system
	Security methodology	Environment variables (.env) used for sensitive credentials (API keys) to prevent unauthorized access.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Privacy depends on the input dataset used by the end-user.
	Intellectual property rights responsible	Jozef Stefan Institute (Dept. for Artificial Intelligence).
	Intellectual property rights constraints	Licensed under MIT License. Open for re-use with attribution.
Access & Sharing	**Receiver(s)**	Data Scientists, Causal Inference Researchers, Developers using StreamStory.
	Re-use methodology	Publicly available code. Users clone via Git and install via uv package manager.
	Any other existing data (asset) re-use	N/A
	Access, sharing & licensing methodology	Hosted on GitHub public repository. Distributed under MIT License.
Archiving & Providing Access	**Preservation planning**	Hosted on GitHub
	Long-term access provision planning	Git history ensures access to previous versions.
	Preservation duration	Indefinite
Respon sibility	**Responsible for data asset management**	Department for Artificial Intelligence, Jozef Stefan Institute; Contributor: Jan Šturm

Table 50 – StreamStory XAI Dashboard

StreamStory XAI Dashboard		
Data Asset Overview	**Description**	An advanced pattern mining and clustering dashboard analyzing hierarchical Markov chain models. It features SHAP-based AI explanations and LLM integration for industrial system insights.
	Source/Origin	GitHub Repository: FAME-JSI/XAI-Dashboard. Developed by the Department for Artificial Intelligence at the Jozef Stefan Institute.
	Collection/Generation methodology	N/A
	Type	Software Application / Visualization Dashboard
	Format	Python (.py), HTML/JS (Frontend), JSON (Input models/LLM outputs), TXT (Descriptions).
	Behavior	Dynamic: Software is updated via commits. Dashboard content changes based on loaded experiment data.
	Expected size	Small. Input Data: Depends on the size of the StreamStory models loaded.
Metadata	**Metadata & Standards used**	pyproject.toml and uv.lock for dependency management. JSON schema for model inputs.
	Documentation methodology	README.md for setup
Storage, Backup & ecurity	**Storage methodology**	Source code on GitHub. Input data stored locally in data/input/ directory structure.
	Backup methodology	Git version control.
	Security methodology	API keys managed via .env file to ensure secure access to external services
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Processes industrial data patterns. Privacy is managed by keeping specific experiment data local.
	Intellectual property rights responsible	Jozef Stefan Institute (Dept. for Artificial Intelligence)
	Intellectual property rights constraints	Licensed under MIT License
Access & Sharing	**Receiver(s)**	Industrial System Operators, Data Analysts, XAI Researchers
	Re-use methodology	Open source. Requires uv for installation and StreamStory models for input
	Any other existing data (asset) re-use	Re-uses pre-generated StreamStory models using StreamStoryPyClient

	Access, sharing & licensing methodology	Public GitHub repository. MIT License
Archiving & Providing Access	**Preservation planning**	Persistent hosting on GitHub
	Long-term access provision planning	Version control history
	Preservation duration	Indefinite
Respon sibility	**Responsible for data asset management**	Department for Artificial Intelligence, Jozef Stefan Institute; Contributor: Jan Šturm

Table 51 – StreamStory Python Client

StreamStory Python Client		
Data Asset Overview	**Description**	A Python client library for creating StreamStory models
	Source/Origin	GitHub Repository: E3-JSI/StreamStoryPyClient. Developed by the Department for Artificial Intelligence at the Jozef Stefan Institute.
	Collection/Generation methodology	Acts as an interface (wrapper) to the StreamStory API.
	Type	Software library
	Format	Python (.py), YAML (examples).
	Behavior	Dynamic: regular releases
	Expected size	Small
Metadata	**Metadata & Standards used**	pyproject.toml for packaging standards
	Documentation methodology	README.md and a comprehensive examples/ directory showing usage patterns.
Storage, Backup & Security	**Storage methodology**	GitHub (Source), PyPI (Package distribution)
	Backup methodology	Git version control.
	Security methodology	Client requires API keys for authentication via .env file
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	The client handles data transmission to the API; privacy depends on the API implementation and user data handling
	Intellectual property rights responsible	Jozef Stefan Institute (Dept. for Artificial Intelligence)
	Intellectual property rights constraints	Licensed under MIT License

Access & Sharing	**Receiver(s)**	Python Developers, IoT Engineers, Data Scientists integrating StreamStory
	Re-use methodology	Installable via pip or uv. Designed for integration into other software stacks
	Any other existing data (asset) re-use	/
	Access, sharing & licensing methodology	Public GitHub, PyPI package registry. MIT License
Archiving & Providing Access	**Preservation planning**	Hosted on GitHub and Python Package Index (PyPI)
	Long-term access provision planning	Release history available on GitHub and PyPI
	Preservation duration	Indefinite
Respon sibility	**Responsible for data asset management**	Department for Artificial Intelligence, Jozef Stefan Institute; Contributor: Jan Šturm

Table 52 – Incremental Analytics

Incremental Analytics		
Data Asset Overview	**Description**	A relational database with capabilities for fast data ingestion and incremental query processing on analytical operations, being energy efficient
	Source/Origin	https://gitlab.gftinnovation.eu/fame/leanxcaledb
	Collection/Generation methodology	N/A
	Type	Software
	Format	Internal data structure of the database
	Behavior	Fixed
	Expected size	Docker image of 2.65GB. Expected storage needs depend on the volume of the data that should be stored
Metadata	**Metadata & Standards used**	Compliant with the JDBC standard for data connectivity
	Documentation methodology	Data lifecycle–based analytics documentation

Storage, Backup & Security	**Storage methodology**	Data are stored within the data nodes of this asset. Permission is granted to authorized users, in a role-based manner, according to the JDBC specifications
	Backup methodology	Periodic backup of data tables can be enabled. The user can define the frequency and the location
	Security methodology	It allows TLS encryption over transmitted data between the database and the client. Data could be also encrypted when stored
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Only certified members within the FAME consortium can access this data asset
	Intellectual property rights responsible	This is under LXS IPR
	Intellectual property rights constraints	This asset cannot be shared or used externally to the FAME consortium without LXS permissions
Access & Sharing	**Receiver(s)**	Data consumers that with to benefit from efficient and incremental query processing that accelerates AI performance. This is pilot agnostic
	Re-use methodology	N/A
	Any other existing data (asset) re-use	N/A
	Access, sharing & licensing methodology	N/A
Archiving & Providing Access	**Preservation planning**	It can be hosted in the project's marketplace. It will be available from the official LXS sales channels
	Long-term access provision planning	Future users may contact LXS directly
	Preservation duration	This is the main product of LXS, it will be preserved as long as the company is working
Respon sibility	**Responsible for data asset management**	Pavlos Kranas LeanXcaleDB pavlos@leanxcale.com

Table 53 – Fast Data Importer

Fast Data Importer		
Data Asset Overview	**Description**	A *side car* process used to load data of huge volume into the Incremental Analytics component, exploiting its ability for fast data ingestion and incremental query processing. It targets IoT datasets for KPI monitoring, containing time-series data

	Source/Origin	https://gitlab.gftinnovation.eu/fame/moh_importer
	Collection/Generation methodology	N/A
	Type	Software
	Format	This do not store any data
	Behavior	Fixed
	Expected size	Docker image of 918MBs
Metadata	**Metadata & Standards used**	N/A
	Documentation methodology	Reproducible data analysis documentation methodology
Storage, Backup & Security	**Storage methodology**	This do not store any data
	Backup methodology	This do not store any data
	Security methodology	It transmits data to the database using TLS encryption
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Only certified members within the FAME consortium can access this data asset
	Intellectual property rights responsible	This is an open-source asset under MIT license
	Intellectual property rights constraints	None
Access & Sharing	**Receiver(s)**	Designed for pilot#7, however it is a generic asset can be applied to any type of time-series dataset that complies with the expected input
	Re-use methodology	They would need to transform their data assets to the expected input of this data asset
	Any other existing data (asset) re-use	N/A
	Access, sharing & licensing methodology	N/A
Archiving & Providing Access	**Preservation planning**	It can be hosted in the project's marketplace.
	Long-term access provision planning	Trusted repositories with persistent identifiers
	Preservation duration	1 year after the end of the project, and this can be extended if there is a need to further maintain this asset
Responsibility	**Responsible for data asset management**	Pavlos Kranas LeanXcaleDB pavlos@leanxcale.com

Table 54 – FML Framework

FML Framework		
Data Asset Overview	**Description**	A framework that enables FML, focusing on the data management aspects. It accompanies the Incremental Analytics and allows the edge nodes to periodically send aggregated data to a central instance. However, even if it fits perfectly with the Incremental Analytics, its scope is generic and it can extend any given relational database
	Source/Origin	https://gitlab.gftinnovation.eu/fame/fl-leanxcale-agent
	Collection/Generation methodology	N/A
	Type	Software
	Format	This do not store any data
	Behavior	Fixed
	Expected size	Docker image of 1.28GB
Metadata	**Metadata & Standards used**	Compliant with the JDBC standard for data connectivity, allowing to be used to any given relational database
	Documentation methodology	N/A
Storage, Backup & Security	**Storage methodology**	This do not store any data
	Backup methodology	This do not store any data
	Security methodology	It transmits data to the databases using TLS encryption
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Only certified members within the FAME consortium can access this data asset
	Intellectual property rights responsible	This is an open-source asset under MIT license
	Intellectual property rights constraints	None
Access & Sharing	**Receiver(s)**	AI analytics that wish to offload the process of syncing the databases in a FML setup
	Re-use methodology	FAIR principles–based data reuse
	Any other existing data (asset) re-use	N/A
	Access, sharing & licensing methodology	N/A

Archiving & Providing Access	**Preservation planning**	It can be hosted in the project's marketplace.
	Long-term access provision planning	Not clear at this phase.
	Preservation duration	Not clear at this phase. The maintenance will be extended if there is a true need for this data asset
Respon sibility	**Responsible for data asset management**	Pavlos Kranas LeanXcaleDB pavlos@leanxcale.com

Table 55 – Marketplace UI Template

Marketplace UI Template		
Data Asset Overview	**Description**	A reusable user interface template for the FAME Data Marketplace, providing a consistent look & feel across all marketplace pages. It includes layout structures, navigation components, forms, cards, typography and styling guidelines, ensuring responsive design, accessibility, and alignment with modern UI/UX practices. The asset is used as the visual and interaction layer for marketplace functionalities (search, browsing, asset details, user account pages, etc.).
	Source/Origin	Internally developed within the FAME project by the Marketplace frontend development team (FAME consortium technical partners). The primary source is the project's code repository via Gitlab.
	Collection/Generation methodology	The template is generated through an iterative software development process: • Requirements gathered from marketplace use cases and WP deliverables. • UX/UI design based on best practices and accessibility guidelines (e.g. WCAG). • Implementation using modern web technologies (HTML5, CSS/SCSS, JavaScript/TypeScript and a contemporary web framework). • Continuous integration, code review and user feedback from pilots.
	Type	APPLICATION
	Format	• Source code files: .ts, .js, .html, .scss / .css • Configuration files: .json, .yaml / .yml

		• Static assets: .svg, .png, .jpg • Build artefacts: bundled/minified web assets (e.g. in dist/ folder)
	Behavior	Changing over time (dynamic)
	Expected size	Overall size (source + assets + build artefacts) is expected to remain below 500 MB.
Metadata	**Metadata & Standards used**	• Standard code metadata (package descriptors such as package.json or equivalent). • Internal documentation of components (e.g. Storybook-style descriptions or markdown docs). • Compliance with web standards: HTML5, CSS3, WAI/WCAG accessibility guidelines.
	Documentation methodology	Developer-oriented documentation (README, component usage guides, architectural overview) stored in the same repository as the code.
Storage, Backup & Security	**Storage methodology**	• Primary storage in a secure, version-controlled Git repository (Gitlab) • Access restricted to authorised FAME consortium members (role-based permissions). • Deployed versions stored on the FAME Marketplace application servers.
	Backup methodology	Server-level backups of deployed instances as part of the project's infrastructure backup plan via Kubernetes and ArgoCD.
	Security methodology	Audit trails/logging available from the Git platform (commit history, access logs).
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	The UI Template does not contain personal or sensitive data; it is purely software and design assets. Privacy impact is minimal. Any potential secrets (e.g. API keys) are excluded from the repository and handled via secure configuration management.
	Intellectual property rights responsible	FAME Consortium, with lead responsibility assigned to the partner(s) responsible for the Marketplace frontend development. Ownership and exploitation are governed by the GA and CA.
	Intellectual property rights constraints	During the project: internal use within the consortium under agreed conditions.

Access & Sharing	**Receiver(s)**	• FAME consortium developers and integrators. • Pilot/application partners building or extending marketplace pages. • Potential external adopters of the FAME Marketplace.
	Re-use methodology	Developers copy or import components from the template into new or existing pages.
	Any other existing data (asset) re-use	Re-uses open-source UI frameworks, CSS libraries and icon packs (e.g. component libraries, typography systems, icon sets), each under its own license.
	Access, sharing & licensing methodology	Internal access via the project's private Git repository with controlled permissions.
Archiving & Providing Access	**Preservation planning**	Final stable release of the UI template will be archived as a tagged version within the repository.
	Long-term access provision planning	N/A
	Preservation duration	As long as it is defined in the exploitation strategy
Respon sibility	**Responsible for data asset management**	Argyro Mavrogiorgou (UPRC), Thanos Kiourtis (UPRC), Nikolaos Zafeiropoulos (UPRC)

Table 56 – Marketplace Chatbot

Marketplace Chatbot		
Data Asset Overview	**Description**	An AI-powered chatbot service integrated into the FAME Marketplace, built on webagent.ai and tailored to FAME content. It provides conversational assistance for navigating the marketplace, discovering assets, and accessing documentation. The asset includes the chatbot backend, integration logic, configuration, and the content used to answer questions (e.g. FAQs and knowledge base material derived from FAME documentation).
	Source/Origin	Developed within the FAME project using the webagent.ai framework and FAME Marketplace documentation.
	Collection/Generation methodology	Configuration of webagent.ai with domain-specific prompts, intents and knowledge sources (e.g. project documentation, FAQs).
	Type	SERVICE
	Format	Source code and integration logic: .py, .js, .ts Configuration & prompts: .json, .yaml / .yml

		Knowledge base: structured and semi-structured text (e.g. .md, .pdf, .html).
	Behavior	Changing over time (dynamic)
	Expected size	Code + configuration: typically, < 100 MB.
Metadata	**Metadata & Standards used**	No
	Documentation methodology	No
Storage, Backup & Security	**Storage methodology**	Source code and configs stored in secure Git repositories under role-based access.
	Backup methodology	N/A
	Security methodology	All communications with the chatbot service are encrypted via HTTPS.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	Data minimisation and limited retention periods are applied for logs.
	Intellectual property rights responsible	FAME Consortium, with lead responsibility assigned to the partner(s) responsible for the Marketplace frontend development. Ownership and exploitation are governed by the GA and CA.
	Intellectual property rights constraints	During the project: internal use within the consortium under agreed conditions.
Access & Sharing	**Receiver(s)**	End-users of the FAME Marketplace (researchers, industry partners, other stakeholders).
	Re-use methodology	The chatbot service is consumed via the Marketplace UI (embedded widget) and/or via APIs.
	Any other existing data (asset) re-use	May re-use external FAQs or public documentation about the FAME project, where permitted.
	Access, sharing & licensing methodology	No
Archiving & Providing Access	**Preservation planning**	Final stable version of the chatbot configuration and integration code will be archived in the repository and/or an institutional archive.
	Long-term access provision planning	N/A
	Preservation duration	As long as it is defined in the exploitation strategy
Respon sibility	**Responsible for data asset management**	Argyro Mavrogiorgou (UPRC), Thanos Kiourtis (UPRC), Nikolaos Zafeiropoulos (UPRC)

Table 57 – FAME Marketplace Information Package

FAME Marketplace Information Package		
Data Asset Overview	**Description**	A comprehensive documentation package for the FAME Data Marketplace, including architecture descriptions, user and administrator manuals, asset publishing guidelines, data discovery workflows, integration and exploitation scenarios, and best practices. It targets both technical and non-technical audiences to support adoption, operation, and exploitation of the Marketplace.
	Source/Origin	Authored within the FAME project by relevant WPs and partners.
	Collection/Generation methodology	Collection of requirements and experiences from pilots, technical development and stakeholder engagement.
	Type	LIBRARY
	Format	Text documents: .docx, .pdf, .odt, .md.
	Behavior	Changing during the project (draft versions, updates), then stabilised in final form at or near project end.
	Expected size	Total size expected to be < 50 MB (depending on number of documents, images and versions).
Metadata	**Metadata & Standards used**	No
	Documentation methodology	Structured sections (introduction, technical documentation, user guides, exploitation scenarios).
Storage, Backup & Security	**Storage methodology**	Final versions stored as part of official project deliverables and potentially mirrored on the FAME marketplace website.
	Backup methodology	No
	Security methodology	Draft versions accessible only to authorised consortium members.
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	The documentation generally does not include personal data.
	Intellectual property rights responsible	FAME Consortium, with lead responsibility assigned to the partner(s) responsible for the Marketplace frontend development. Ownership and exploitation are governed by the GA and CA.
	Intellectual property rights constraints	During the project: internal use within the consortium under agreed conditions.

Access & Sharing	**Receiver(s)**	• FAME consortium members (developers, administrators, pilots). • External users and adopters of the FAME Marketplace (technical and non-technical).
	Re-use methodology	No
	Any other existing data (asset) re-use	No
	Access, sharing & licensing methodology	FAIR-compliant open access data sharing
Archiving & Providing Access	**Preservation planning**	As long as it is defined in the exploitation strategy
	Long-term access provision planning	N/A
	Preservation duration	As long as it is defined in the exploitation strategy
Respon sibility	**Responsible for data asset management**	Argyro Mavrogiorgou (UPRC), Thanos Kiourtis (UPRC), Nikolaos Zafeiropoulos (UPRC)

Table 58 – Anomaly Detection FL Center Server

Anomaly Detection FL Center Server		
Data Asset Overview	**Description**	This asset consists of a Docker container acting as the center server in hierarchical deployments for the training of an anomaly detector neural network in a Federated Learning manner.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Ffederated_learning%2Fcenter_server/artifacts-tabarbor
	Collection/Generation methodology	A software tool built upon ATOS proprietary Federated Learning Framework
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	3.01GB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
ge, Backu p &	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup

	Security methodology	Access control, code obfuscation
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS, according to the conditions established in the Consortium Agreement.
Access & Sharing	**Receiver(s)**	Pilot 7 owners and stakeholders interested in Federated Learning for industrial environments
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source. Pilot 7 data was only used to evaluate the implementation as a proof-of-concept.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by the asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 59 – Anomaly Detection FL Client

Anomaly Detection FL Client		
Data Asset Overview	**Description**	This asset consists of a Docker container for the training of an anomaly detector neural network in a Federated Learning manner.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Ffederated_learning%2Fclient/artifacts-tab
	Collection/Generation methodology	A software tool built upon ATOS proprietary Federated Learning Framework
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	3.01GB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control, code obfuscation
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS, according to the conditions established in the Consortium Agreement.
Access & Sharing	**Receiver(s)**	Pilot 7 owners and stakeholders interested in Federated Learning for industrial environments
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source. Pilot 7 data was only used to evaluate the implementation as a proof-of-concept.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.

Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 60 – Anomaly Detection FL Root Server

Anomaly Detection FL Root Server		
Data Asset Overview	**Description**	The Anomaly Detection FL Root Server asset consists of a Docker container aggregating weights and sending them back to the clients.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Ffederated_learning%2Froot_server/artifacts-tab
	Collection/Generation methodology	A software tool built upon ATOS proprietary Federated Learning Framework
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	3.01GB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control, code obfuscation

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS, according to the conditions established in the Consortium Agreement.
Access & Sharing	**Receiver(s)**	Pilot 7 owners and stakeholders interested in Federated Learning for industrial environments
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source. Pilot 7 data was only used to evaluate the implementation as a proof-of-concept.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 61 – Federated KMeans Server for Financial Services

Federated KMeans Server for Financial Services		
Data Asset Overview	**Description**	The Federated KMeans Server asset is a Docker container aggregating KMeans centroids and sending the aggregated centroids back to its clients via HTTP.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Ffederated_learning%2Fkmeans_server/artifacts-tab
	Collection/Generation methodology	A software tool built upon ATOS proprietary Federated Learning Framework
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	881.53MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Pilot 1 owners and stakeholders interested in Federated Learning for the banking sector
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source. Pilot 1 data was only used to evaluate the implementation as a proof-of-concept.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.

Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 62 – Federated KMeans Client for Financial Services

Federated KMeans Client for Financial Services		
Data Asset Overview	**Description**	The Federated KMeans Client asset is a Docker container training KMeans models and sending the model to its server via HTTP.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Ffederated_learning%2Fkmeans_client/artifacts-tab
	Collection/Generation methodology	A software tool built upon ATOS proprietary Federated Learning Framework
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	881.53MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Pilot 1 owners and stakeholders interested in Federated Learning for the banking sector
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source. Pilot 1 data was only used to evaluate the implementation as a proof-of-concept.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 63 – SD API-Server

SD API-Server		
Data Asset Overview	**Description**	Application to build everything needed to deploy the models and connects with the rest of the Smart Deployment applications.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/smartdeployment%2Fapi-server/artifacts-tab
	Collection/Generation methodology	The data asset implements a RESTful API to trigger its methods.
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	824.61MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Stakeholders interested in Smart Deployment for their services deployment.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.

Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 64 – SD Node-Resolver

SD Node-Resolver		
Data Asset Overview	**Description**	Application to decide in which node the model must be deployed due to the CO_2 emissions from the country node.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/smartdeployment%2Fnode-resolver/artifacts-tab
	Collection/Generation methodology	The data asset implements a RESTful API with the methodology of the node election.
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	787.26MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Stakeholders interested in Smart Deployment for their services deployment.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 65 – SD Model-Server

SD Model-Server		
Data Asset Overview	**Description**	Application to serve the model.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/smartdeployment%2Fmodel-server/artifacts-tab
	Collection/Generation methodology	The data asset is a RESTful API with the functionality to serve a model

	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	2.01GB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Stakeholders interested in Smart Deployment for their services deployment.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions

Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 66 – SD WebApp

SD WebApp		
Data Asset Overview	**Description**	Application to give to the user an user-friendly webapp to use the Smart Deployment API-Server to deploy the models.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/smartdeployment%2Fwebapp/artifacts-tab
	Collection/Generation methodology	The data asset implements a UI that integrates with the rest of components of the Smart Deployment to communicate with the user
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	568.18MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Stakeholders interested in Smart Deployment for their services deployment.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed

		according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any data source.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 67 – LSTM Forecasting

LSTM Forecasting		
Data Asset Overview	**Description**	Time-series forecasting model for machinery temperature sensors.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fpilot7_forecasting/artifacts-tab
	Collection/Generation methodology	The data asset implements a RESTful API to trigger model inference and obtain the result using a Long Short-Term Memory (LSTM) model.
	Type	Model
	Format	Docker container
	Behavior	Fixed
	Expected size	806.17MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description

Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Pilot 7 owners and stakeholders interested in Smart Deployment for their services deployment.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The model was trained with data from Pilot 7.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 68 – TabNet Classifier Model

TabNet Classifier Model		
Data Asset Overview	**Description**	TabNet model for default risk prediction on a financial service
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Ftabnet_classifier_model/artifacts-tab
	Collection/Generation methodology	The fraud prediction model identifies customers who may fall behind on payments. The training process incorporated SMOTE resampling, as the dataset was highly imbalanced and contained only 2.6% positive instances.
	Type	Model
	Format	Docker container
	Behavior	Fixed
	Expected size	351.59MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Internal risk management and credit assessment teams within the financial institution.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The model was trained with data from Pilot 1.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted

		by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 69 – K-Means Model

K-Means Model		
Data Asset Overview	**Description**	Customer segmentation using K-Means implemented in a financial services context
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fkmeans_financial_service_model/artifacts-tab
	Collection/Generation methodology	A K-Means clustering model was implemented using the scikit-learn[1] library, resulting in four customer segments. The dataset was processed using the Recency, Frequency and Monetary (RFM) technique. The model was evaluated using a five-fold cross-validation strategy with shuffled splits.
	Type	Model
	Format	Docker container
	Behavior	Fixed
	Expected size	68.62MB

[1] https://scikit-learn.org/

Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Internal business teams focused on customer segmentation insights for strategic decision-making.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The model was trained with data from Pilot 1
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 70 – K-Means Model

K-Means Model		
Data Asset Overview	**Description**	Customer segmentation using K-Means implemented in a financial services context
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fkmeans_financial_service_model/artifacts-tab
	Collection/Generation methodology	A K-Means clustering model was implemented using the scikit-learn[2] library, resulting in four customer segments. The dataset was processed using the Recency, Frequency and Monetary (RFM) technique. The model was evaluated using a five-fold cross-validation strategy with shuffled splits.
	Type	Model
	Format	Docker container
	Behavior	Fixed
	Expected size	68.62MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Internal business teams focused on customer segmentation insights for strategic decision-making.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.

[2] https://scikit-learn.org/

	Any other existing data (asset) re-use	The model was trained with data from Pilot 1
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 71 – K-Means Model

K-Means Model		
Data Asset Overview	**Description**	Customer segmentation model implemented using K-Means for parking usage in Athens
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fkmeans_parking_athens_model/artifacts-tab
	Collection/Generation methodology	A K-Means clustering model was implemented using the scikit-learn library, resulting in four customer segments. The dataset was processed using the Recency, Frequency and Monetary (RFM) technique. The model was evaluated using a five-fold cross-validation strategy with shuffled splits.
	Type	Model
	Format	Docker container
	Behavior	Fixed
	Expected size	68.62MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description

Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Mobility and urban planning team focused on customer segmentation insights for strategic decision-making.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The model was trained with data from Pilot 2
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 72 – LSTM Model

LSTM Model		
Data Asset Overview	**Description**	Large language model trained to extract variables from CDTI documents.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fllm_flan_t5_small_model/artifacts-tab
	Collection/Generation methodology	A FLAN-tuned T5-small sequence-to-sequence language model with a LoRA weights adapter, trained to extract specific variables from the text of CDTI documents.
	Type	Model
	Format	Docker container
	Behavior	Fixed
	Expected size	293.93MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Internal data processing and document analysis teams who need automated extraction of variables from CDTI documents.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The model was trained with data from Pilot 4.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.

Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 73 – RFM Processing Pipeline

RFM Processing Pipeline		
Data Asset Overview	**Description**	Pipeline for computing RFM metrics (recency, frequency, monetary) from customer datasets.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Frfm_processing_pipeline/artifacts-tab
	Collection/Generation methodology	A pipeline designed to calculate the Recency, Frequency, and Monetary (RFM) metrics from datasets containing records of customer movements across any domain.
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	864.43MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control

Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Data engineering and analytics teams for customer insights.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The pipeline was tested using data from Pilot 1 and 2.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 74 – RFM Processing Pipeline

RFM Processing Pipeline		
Data Asset Overview	**Description**	Pipeline for extracting program details and financial metrics from CDTI documentation
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Finformation_extractor/artifacts-tab

	Collection/Generation methodology	A pipeline designed to extract relevant information from official CDTI documents. It requires the PDF file and the specification of the funding category as inputs. The pipeline first performs PDF data extraction using pdfplumber[3] and Camelot[4], and subsequently applies a pre-trained LLM model (asset number 15) to obtain the targeted numerical values. The results are returned in JSON format.
	Type	Software
	Format	Docker container
	Behavior	Fixed
	Expected size	4.3GB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Document analysis teams who need automated extraction of variables from CDTI documents.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The software is standalone and can operate with any new version of CDTI official funding documentation.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted

[3] https://pypi.org/project/pdfplumber/
[4] https://pypi.org/project/camelot-py/

		by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 75 – Parking Zones Dataset

Parking Zones Dataset		
Data Asset Overview	**Description**	Analysis of parking customers in Athens based on customer segmentation results.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fdataset_pilot2_parking_zones/artifacts-tab
	Collection/Generation methodology	Dataset containing six months of data related to parking transactions in Athens. Based on a Recency, Frequency, and Monetary (RFM) dataset, customer segmentation was applied to classify customers into four groups. The most frequently used parking zones were analysed per customer and reported in the dataset.
	Type	Dataset
	Format	CSV file encapsulated in a Docker container
	Behavior	Fixed
	Expected size	378.27MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description

Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Mobility planning and parking management teams in Athens
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The analysis was performed using the dataset from Pilot 2.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 76 – RFM Dataset Pilot 2

RFM Dataset Pilot 2		
Data Asset Overview	**Description**	Recency, Frequency and Monetary (RFM) dataset containing data from parking transactions in Athens.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fdataset_pilot2_rfm/artifacts-tab
	Collection/Generation methodology	This dataset contains six months of data related to parking transactions in Athens. It includes the Recency, Frequency, and Monetary (RFM) metrics for each customer. - The recency metric indicates the number of days since the last parking transaction. - The frequency metric shows how often the customer makes transactions. - The monetary metric reflects the total amount spent by each customer.
	Type	Dataset
	Format	CSV file encapsulated in a Docker container
	Behavior	Fixed
	Expected size	378.23MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Pilot 2 owners and stakeholders interested in Recency-Frequency-Monetary customer analysis
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the

		Consortium Agreement.
	Any other existing data (asset) re-use	This dataset was generated using Pilot 2 data and the RFM pipeline asset (asset 16 above)
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 77 – RFM Dataset Pilot 1

RFM Dataset Pilot 1		
Data Asset Overview	**Description**	Recency, Frequency and Monetary (RFM) dataset containing data of embedding financial services.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fdataset_pilot1_rfm/artifacts-tab
	Collection/Generation methodology	This dataset for embedding financial services contains one year of historical customer data. It includes the Recency, Frequency, and Monetary (RFM) metrics for each customer. - The recency metric indicates how recently the last customer interaction with the service occurred. - The frequency metric shows how often interactions occur. - The monetary metric reflects the amount spent by each customer.
	Type	Dataset
	Format	CSV file encapsulated in a Docker container

	Behavior	Fixed
	Expected size	379.06MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Pilot 1 owners and stakeholders interested in Recency-Frequency-Monetary customer analysis
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	This dataset was generated using Pilot 1 data and the RFM pipeline asset (asset 14 above)
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions

Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 78 – Pilot 4 Synthetic Dataset

Pilot 4 Synthetic Dataset		
Data Asset Overview	**Description**	Synthetic dataset reflecting the structure and content of CDTI documents
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fdataset_pilot4_synthetic_dataset/artifacts-tab
	Collection/Generation methodology	Synthetic dataset containing extracted sections from CDTI documentation. Each input sample consists of text extracted from the official PDFs with relevant numerical variables that usually appear in this documentation. Paraphrasing has been applied in order to augment the dataset with more robust data. Output samples contain the numerical values the model is expected to extract from the text.
	Type	Dataset
	Format	CSV file encapsulated in a Docker container
	Behavior	Fixed
	Expected size	378.06MB
Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS

Access & Sharing	**Receiver(s)**	Pilot 4 owners.
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	The synthetic dataset was generated using Pilot 4 original dataset.
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.
Archiving & Providing Access	**Preservation planning**	The asset is stored in the infrastructure that supports the marketplace beyond the project duration.
	Long-term access provision planning	Access can be granted to new users beyond the duration of the project while the marketplace is active. Internal users will be granted access as per the Consortium Agreement. For external users, access conditions will be negotiated on a case-by-case basis.
	Preservation duration	4 years, as established in the Horizon Europe conditions
Responsibility	**Responsible for data asset management**	Pablo Gil ATOS pablo.gil@eviden.com Raquel Lazcano ATOS raquel.lazcano@eviden.com

Table 79 – Sensor Anomaly Detector

Sensor Anomaly Detector		
Data Asset Overview	**Description**	Autoencoder that detects anomalies in machinery temperature sensor.
	Source/Origin	https://harbor.gftinnovation.eu/harbor/projects/40/repositories/ai_tools%2Fai_analytics%2Fsensor_anomaly_detector/artifacts-tab
	Collection/Generation methodology	An autoencoder trained to detect statistical anomalies from industrial sensor data.
	Type	Model
	Format	TensorFlow model encapsulated in Docker container
	Behavior	Fixed
	Expected size	1.48GB

Metadata	**Metadata & Standards used**	It does not apply to this asset.
	Documentation methodology	README included in the asset description
Storage, Backup & Security	**Storage methodology**	Project's Harbor Sever
	Backup methodology	Harbor's backup
	Security methodology	Access control
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	ATOS Proprietary, privacy & confidentiality should be granted by the marketplace
	Intellectual property rights responsible	ATOS Proprietary
	Intellectual property rights constraints	Permission should be granted by ATOS
Access & Sharing	**Receiver(s)**	Pilot 7 owners and stakeholders interested in sensor anomaly detection in industrial environments
	Re-use methodology	External stakeholders must request access to asset owner. Access conditions will be negotiated on a case-by-case basis. For FAME partners, this access will be managed according to the conditions established in the Consortium Agreement.
	Any other existing data (asset) re-use	Pilot 7 data was used to train the model
	Access, sharing & licensing methodology	No licensing applies as the asset ownership remains proprietary. Access must be granted by asset owner.

4 FAIR Data & Zenodo Platform in FAME

4.1 FAIR Data

The Federated Authentication and Authorization Infrastructure (AAI) developed in *T3.1 - Federated AAI Infrastructure* implements an infrastructure over different types of data providers and their data infrastructures (i.e., data marketplaces, data spaces, databases etc.). The project leverages the existing i3-MARKET platform, which provides a baseline support for self-sovereign identities and access to data from federated marketplaces. The platform is extended with interfaces that adhere to the needs of FAME. Specifically, interfaces for data assets trading, pricing and data policy management are implemented, while also, existing interfaces are enhanced to support additional data exchange models and ontologies. FAME ontologies and models for EmFi are specified in *T3.4 - Semantic Interoperability Middleware*.

To secure findability and data reuse, a federated catalogue that comprises data assets from the various federated data sources and marketplaces is introduced/extended in *T3.3 - Federated Catalogue of Data Assets*. In this direction, the various data assets are modelled according to the FAME models and are structured according to EmFi ontologies and taxonomies. The Federated Catalogue of Data Assets also specifies and implements APIs for accessing, searching, and querying the federated catalogue. These APIs ensure a seamless and marketplace agnostic experience for the users of FAME in terms that the users will be able to discover data assets across different marketplaces and data spaces.

The management of data access policies in the FAME Data Marketplace is handled in *T3.2 - Unified Security Policy Management*. The implementation of a security policies management tool aims to map FAME policies to the lower-level policies of the underlying providers, aligning to the FAIR principles. The tool enables the specification of policies over the federated marketplaces. The starting point for the specifications is already existing ontologies of the finance sector (e.g., FIBO, FIGI, H2020 INFINITECH ontologies). These ontologies will be linked with ontologies from other sectors (e.g., retail, smart cities, healthcare) in-line with the requirements of EmFi use cases. Moreover, the task implements a middleware that will transform data from the formats and semantics of the individual underlying marketplaces and data spaces to the FAME ontologies. In this direction, the task leverages the AAI interfaces to access the data and annotate semantically the assets.

Furthermore, search engine results ranking techniques to implement dynamic market mechanisms for data assets trading and monetization are leveraged in the scope of *T4.4 - Semantic Search for Trading and Valuation of Data Assets*. In more detail, this task implements semantic search over the federated catalogue along with schemes for ranking the results according to relevance and value-based attributes of the data assets and is an important component in the overall data management of FAME and its federated data assets.

A wide range of data cleaning and interoperability standards-based techniques are employed to detect and correct corrupt or inaccurate records and homogenize the collected data (complementing the functionalities of *T3.4 – Semantic Interoperability Middleware* for Interoperability).

Below, after a short presentation of each FAIR relevant sub-principle based on the “FAIR Guiding Principles for scientific data management and stewardship” and the “Three-point FAIRification

Framework", the corresponding Tasks that indicatively implements the referred mechanisms in compliance with FAIR principles are summarised in Table 80.

Table 80 – FAME FAIR Sub-Principles

FAIR Sub-Principles	FAME Tasks
F1. (Meta)data are assigned a globally unique and persistent identifier	T3.3, T3.4, T4.1
F2. Data are described with rich metadata	T3.3, T3.4, T4.1
F3. Metadata clearly and explicitly include the identifier of the data they describe	T3.3, T3.4, T4.1
F4. (Meta)data are registered or indexed in a searchable resource	T3.3
A1. (Meta)data are retrievable by their identifier using a standardized communication protocol	T3.1, T3.3, T3.4, T4.4
A1.1 The protocol is open, free, and universally implementable	T2.4, T3.1, T3.3
A1.2 The protocol allows for an authentication and authorization procedure, where necessary	T2.4, T3.1
A2. Metadata are accessible, even when the data are no longer available	T3.1, T3.3, T3.4, T4.4
I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	T3.4
I2. (Meta)data use vocabularies that follow FAIR principles	T3.2, T3.4
I3. (Meta)data include qualified references to other (meta)data	T3.3, T3.4
R1. (Meta)data are richly described with a plurality of accurate and relevant attributes	T3.2, T3.3, T4.1
R1.1. (Meta)data are released with a clear and accessible data usage license	T3.2, T3.5
R1.2. (Meta)data are associated with detailed provenance	T3.4, T4.1
R1.3. (Meta)data meet domain-relevant community standards	T3.4

4.1.1 Making data findable, including provision for metadata

4.1.1.1 F1. (Meta)data are assigned a globally unique and persistent identifier

This sub-principle refers to the uniqueness and persistence of the identifiers that FAME develops. To ensure the findability of data and research outputs of the project, globally unique and persistent identifiers (PID) are provided (e.g., DOIs).

4.1.1.2 F2. Data are described with rich metadata

This sub-principle dictates that metadata shall be extensive, with reliable information regarding the data's context, quality, condition and characteristics. Metadata of the project's datasets must be machine-readable to optimize their discovery. A URL connected to a document that contains machine-readable metadata and the format of the file is provided.

4.1.1.3 F3. Metadata clearly and explicitly include the identifier of the data they describe

This sub-principle measures whether the metadata document includes the globally unique and persistent identifier for the digital resource. Because metadata and the datasets they describe are in different files, these two files must be associated by describing a dataset's globally unique and persistent identifier in the metadata file. Metadata shall contain the identifier of the digital resource it describes. As mentioned in the **F1** description, many repositories address these issues. An

example is given though http://go-fair.org regarding this connection annotated in a formal manner using the *foaf:primaryTopic* predicate in the case of RDF metadata.

4.1.1.4 F4. (Meta)data are registered or indexed in a searchable resource

This sub-principle refers to finding digital resources using web-based search engines. Search engines must index the digital resource and its metadata for the users to find and reuse them. This can be provided by the persistent identifiers of the resource and URLs that give search results on various search engines.

4.1.2 Making data accessible

4.1.2.1 A1. (Meta)data are retrievable by their identifier using a standardized communication protocol

4.1.2.1.1 A1.1 The protocol is open, free, and universally implementable

This FAIR criterion examines whether the access protocol is free (no-cost) and open-sourced, facilitating data reuse. These two criteria refer at least to the metadata. The project's datasets are available as open data in various FAIR repositories (e.g., Zenodo).

4.1.2.1.2 A1.2 The protocol allows for an authentication and authorization procedure, where necessary

The conditions under which the data are accessible are defined in this sub-principle. Thus, "closed" or protected data are also FAIR. Furthermore, the specification of a protocol to access restricted content is defined for content that is available under restriction. In cases where datasets are confidential, the project considers authentication and authorization in specific accessing procedures.

4.1.2.2 A2. Metadata are accessible, even when the data are no longer available

This metric examines the existence of metadata even in the absence of data. This principle states that metadata will persist over time, continuing to provide descriptors of the data in connection with ***F3-F4*** principles (metadata remains discoverable with a persistent identifier).

4.1.3 Making data interoperable

4.1.3.1 I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

To ensure the automatic findability and interoperability of datasets, FAME adopts commonly used vocabularies and ontologies and a well-defined data model that describes the structure of the metadata (e.g., RDF knowledge representation model). *T3.4 - Semantic Interoperability Middleware* provides the mechanisms for choosing a suitable schema ensuring schema interoperability through metadata annotation.

4.1.3.2 I2. (Meta)data use vocabularies that follow FAIR principles

Metadata should use community-accepted vocabulary and terminology that are themselves FAIR. The provenance descriptors and vocabulary used within datasets must be documented and resolvable using globally unique and persistent identifiers. FAME incorporates and extends the Mobility Data Specification (MDS) and the outcomes of the Open Mobility Foundation (OMF).

4.1.3.3 I3. (Meta)data include qualified references to other (meta)data

Relationships within the various multidimensional datasets and (meta)data that are used towards the FAME project, have explicit and useful semantic meaning. An example of semantics of these relationships is: "*objectA wasFoundInTheSamePlaceAs objectB*", which is an acceptable qualified reference. This principle addresses the specification of: a) interlinking between datasets, b) the

implementation (if needed) of an additional dataset to complete the data, c) information about data that are stored in a different dataset.

4.1.4 Increase data re-use

4.1.4.1 R1. (Meta)data are richly described with a plurality of accurate and relevant attributes

4.1.4.1.1 R1.1. (Meta)data are released with a clear and accessible data usage license

This sub-principle refers to legal interoperability by defining the conditions under which the metadata can be reused. These conditions are clarified through a license document for data and its associated metadata. This document is provided by the IRI of license (e.g., URL) for both the data and metadata license.

4.1.4.1.2 R1.2. (Meta)data are associated with detailed provenance

This sub-principle ensures that there is provenance information about the data, mainly regarding two topics of provenance: a) Who/what/when produced the data (i.e., whom to cite, who generated or collected the data), b) why/how the data was produced (i.e., description of the workflow that led to the data, to understand its context). This information is delivered by *T3.2 - Unified Security Policy Management*, depicting the overall provenance of the data that is processed during the project, using ontologies and state-of-the-art approaches (e.g., PROV-O ontology). For instance, a record trail that accounts for the origin of a piece of data (in a database, document or repository) together with an explanation of how and why it got to the present place.

4.1.4.1.3 R1.3. (Meta)data meet domain-relevant community standards

Many communities have established metadata guidelines to maximize their shared data's practical usability. As a result, they provide validation services to certify the compliance of the data. Furthermore, these communities offer several standards and requirements regarding context and provenance to improve the reusability of the data. Towards this direction, in the context of FAME, certifications and validation techniques (e.g., a Verisign signature) are implemented. Towards this end, the project already collaborates with various standardization organizations such as GAIA-X and IDSA.

4.2 FAME FAIR Data through the Zenodo Platform

4.2.1 Zenodo Platform in a Nutshell

Zenodo is a widely used open-access platform developed by CERN under the European OpenAIRE program, in May 2013. It enables researchers and EU-funded projects to store, share, and preserve a wide range of research outputs-including publications, datasets, software, and other digital assets-in a secure and sustainable way. Zenodo assigns DOIs to all uploads, ensuring long-term accessibility and proper citation. Additionally, it helps researchers receive credit by making the research results citable and through OpenAIRE integrates them into existing reporting lines to funding agencies like the EC. For FAME, using Zenodo is a strategic step towards making every open data FAIR. By depositing publications and datasets on Zenodo, the visibility and traceability of FAME work, reproducibility, and alignment with the EU's open science principles are enhanced. It ensures that the project's scientific contributions remain accessible to the broader research community, policy makers, and the public, even beyond the project's duration.

4.2.2 Zenodo Platform in FAME

To support open science and ensure all project outputs are FAIR, the FAME project uses Zenodo (https://zenodo.org/communities/fame-project/records?q=&l=list&p=1&s=10&sort=newest), hosting all relevant public-facing content produced within the project, making it freely available to

the research and innovation community. This content is related to publications, public deliverables, or public data assets among others. The FAME Zenodo community is publicly accessible, meaning anyone can view or download uploaded content. The FAME Zenodo structure is organized by the type of the output, including mainly scientific publications, since the related open-access deliverables and other material needs to be reviewed and accepted prior to their publication. All in all, this structure improves accessibility and transparency for both internal stakeholders and external audiences such as reviewers, researchers, and policymakers, supporting also the alignment with HEU's open data mandates.

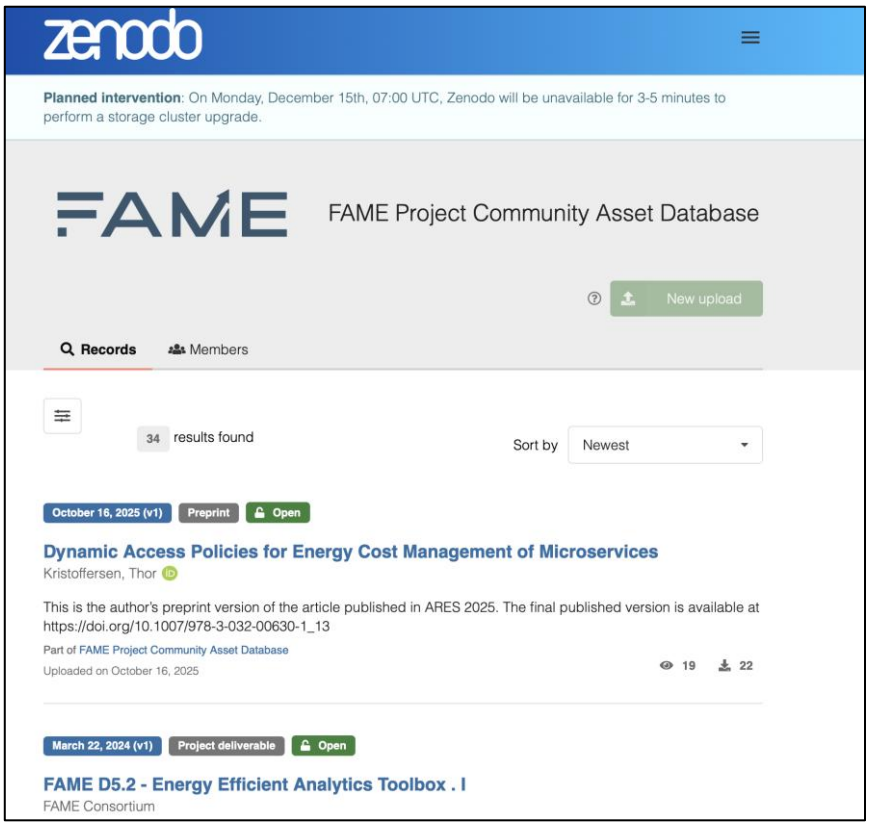

Figure 1 - FAME on Zenodo

Generally, to maintain consistency and ensure dissemination of high-quality, the FAME project follows a specific internal workflow for uploading materials. Each partner first contacts the dissemination team to inform them about the related content that is planned to be published, and accordingly the interested partners need to complete a specific form with information related to their publishment. This form contains essential metadata such as the title and the type of the specific scientific paper. Once validated, the dissemination team handles the uploading to Zenodo, ensuring it is correctly tagged and visible within the FAME community. This coordinated process ensures that the project's research outputs are well-documented, easily discoverable, and aligned with the FAIR principles that govern responsible open science.

4.2.3 Zenodo throughout FAME lifecycle

To ensure that the Zenodo platform remains a dynamic and reliable repository throughout the FAME project's lifecycle, it is planned to regularly enrich Zenodo with key public-facing outputs. These include all public deliverables, scientific publications, relevant dissemination materials, and open-access data assets related to the project partners and their external collaborations. Each upload is reviewed and coordinated by the FAME dissemination and communication team to maintain quality, consistency, and proper alignment with FAIR principles. Looking ahead, FAME is committed to keeping the Zenodo community active even after the project ends. To achieve this, it is ensured that all final public results are uploaded regularly, even after the project's completion

5 Allocation of Resources

5.1 Costs for Making Data FAIR

The project budget includes the costs necessary to make the data gathered/generated during FAME operations FAIR. These projected expenses are required to pay for several specialized data processing and data management tasks, ranging from data collection and documentation through storage and preservation to sharing and reuse. The required effort is included in the relevant WP as these activities are a part of the WP under which the relevant data are processed. The adoption of Zenodo (Section 4.2) will guarantee that the costs associated with long-term data preservation are reasonable.

5.2 Data Management Responsibilities

The responsibilities related to data management are specified as follows:

- **WP Leaders (WPL)** are in charge of organizing the data processing and quality control tasks carried out under the WP they are directing.
- **Task Leaders (TL)** operate as Data Controllers for the data gathered or created throughout the tasks they are in charge of. Additionally, they ensure that the data are correctly packaged for sharing with partners or for public consumption.
- The **consortium partners** who process the data created or collected are referred to as *Data Processors*. These processes involve data gathering, digitization, anonymization, storage, deletion, quality assurance, and consent acquisition.
- **Data Repositories** are intended for the long-term storage and preservation of project data. In earlier sections of this deliverable, Zenodo (Section 4) and the FAME website [9] were both discussed and described in relation to public and open data. As a result, the project's document repository that is only accessible by authorized project partners must safely store and maintain the data that can be shared among authorized consortium members during the project.

6 Data Security

Research ethics' main concern is data protection. All people have the right to control how information on them is gathered and used, which is protected by the EU Charter of Fundamental Rights. Everyone has the right to the protection of their own personal data under Article 8(1) of the EU's Charter of Fundamental Rights and Article 16(1) of the Treaty on the Functioning of the EU (TFEU), and the GDPR establishes guidelines for the protection of natural persons regarding the processing of personal data. Anywhere that personal data - information pertaining to a named or distinguishable real person - is processed [10], extreme caution must be exercised. Data protection laws require researchers conducting study to give research participants relevant information about what will happen to the personal data they collect. Research involving sensitive data, must receive special consideration because, per GDPR, such data cannot be processed without the subject's explicit consent or one of the other legal bases listed under Article 9(2) of the GDPR, such as processing "necessary for archiving purposes in the public interest scientific or historical research […] [GDPR, Art. 9(2)(j)]". According to the GDPR, the legal bases for processing personal data pertaining to stakeholders are that each data subject employed by a project partner has consented to the processing of his or her personal data (Article 6(1)(a)), and the processing is required for the performance of a contract, in this case, the employment contracts between the data subjects and each project partner (Article 6(1)(b)). Each project partner will keep records including the personal information of those who worked for them after the project is finished. Unless said employees prefer that their contact information be deleted, every partner is free to keep its copy of the contact information of employees who work for other FAME partners. Only after the very last payment and evaluation from the EC, the mailing lists for the project will be removed. Only project participants will have access to the contact information of data subjects who are employees of any of the project partners, and only for as long as it takes to execute the GA and/or finish the project. Once the application is made publicly or commercially available, authorship information may be made public with the consent of the data subjects who are the authors.

All partners who process data have a responsibility to keep it secure by implementing all essential access controls (identification, authentication, and authorization) as well as security controls (including backup procedures and integrity checks) inside their infrastructure. In the unfortunate event of a personal data breach, the project partners will promptly notify both the data subject(s) who may be impacted by the incident and their respective national supervisory authorities. At the same time, they will record any breaches of personal data and any relevant details. If a data breach poses a risk to data subjects, a data controller must notify the supervisory authority within 72 hours of becoming aware of it [11]. Any breach of personal data must be documented by the controller, describing the circumstances surrounding the breach, its effects, and the steps that were taken to address them.

In FAME, regulatory compliance is viewed as a horizontal duty that affects all consortium members. All project participants are expected to make sure that the necessary steps are taken to prevent and lessen the effects of any incidents involving the security of the platform, the network, and the information systems. To ensure that the project's operations continue in a safe manner, all participants should execute an appropriate continuity strategy. To minimize negative impacts and stop similar incidents from happening in the future, it is important to evaluate the impact of each breach of personal information. When determining the significance of a personal data breach, several variables need to be considered, such as the quantity of individuals impacted, the severity of the impact on their rights and freedoms, as well as on project activities, and the length of the occurrence.

Each pilot partner already has a Data Protection Officer (DPO), where in case of a potential data breach, mitigation actions are considered in accordance with all the DPOs.

7 Ethical Aspects

FAME project carries out in line with the highest ethical standards and the applicable EU, international and national law on ethical principles. Ethics are considered significant in FAME project and are a distinct part of the GA, while the ethical requirements are one of the purposes of WP1, whose objective is to ensure ethical compliance with the ethical requirements' set.

This reflects to the report of the ethical standards of the FAME project, which develops the project's social, legal, and ethical activities framework, including the project's methodology for handling social science and gender issues. The framework ensures the compliance of the project's technologies to applicable laws and regulations (e.g., 4AML, PSD2, MiFIDII, GDPR), as well as to emerging regulations (e.g., AI regulation Act). It also deals with issues like data sources specification and co-creation, identification of the environmental dimension of the research, and social impact assessment. Moreover, the task identifies ethics and risks associated with the project's technologies, along with guidelines to mitigate them. For the work of the task to be performed and to establish a solid ethical framework, the task leader appointed an Ethics and Legal Expert with knowhow on financial regulations. It also mobilizes the FAME Ethical Board (EB), which audits the research activities against ethical and regulatory compliance. The designed ethical framework also includes frequent interactions with the use cases towards ensuring their compliance to regulatory requirements based on the FAME technologies and tools.

The EB continuously provides oversight, monitoring progress and addressing any ethical challenges that may emerge. This process reinforces the project's mission to implement tools that are safe, transparent, and ethically sound, offering meaningful and trustworthy solutions in the EmFi domain.

8 Conclusions

This final version of the FAME DMP presented the consolidated data management framework adopted throughout the project, fully aligned with open science and open data requirements. It defined the principles, processes, and responsibilities governing the management of data assets generated, collected, and exchanged within FAME and the FAME Data Marketplace, with a particular emphasis on maximising data FAIRness. Since in this version of the DMP focus was given to the data assets already indexed in the FAME Data Marketplace, it should be mentioned that for the ones that have not been indexed yet, a, identical approach will be followed for all of them.

The FAME DMP is grounded in established technical approaches, tools, and standards for research data management and data sharing. Trusted repositories and catalogues, including OpenAIRE-compliant infrastructures such as Zenodo, are leveraged to support the long-term preservation, discoverability, and citation of selected research outputs, publications, and open data assets. In parallel, the FAME Data Marketplace and the project's communication channels provide additional access points to publicly available information, documentation, and dissemination materials, supporting engagement with the wider data economy and research communities. These combined mechanisms ensure that data assets and related outputs remain accessible and usable beyond the lifetime of the project.

At the same time, the DMP places strong emphasis on data protection, confidentiality, and compliance with applicable regulatory frameworks, including GDPR. Appropriate safeguards, governance mechanisms, and informed consent procedures have been defined and applied to ensure that personal, sensitive, or commercially confidential data assets are managed responsibly.

This DMP reflects the mature understanding gained during the implementation of the FAME project regarding the nature, lifecycle, and exploitation potential of its data assets. It documents the practices adopted by project partners and provides a stable reference framework for future reuse, sustainability, and exploitation of FAME results. Through this final DMP, FAME delivers a robust, transparent, and standards-based foundation for responsible data sharing and marketplace-driven data exploitation, in line with Horizon Europe objectives and long-term impact expectations.

References

1. Data Management Plan Template, https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf (retrieved 2025/12/15)
2. Go FAIR - FAIR Principles, https://www.go-fair.org/fair-principles/ (retrieved 2025/12/15)
3. FAME, D1.2 - Data Management Plan I, University of Piraeus Research Centre, (2023).
4. Wilkinson, M., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship, Sci Data3, 160018
5. EOSC Declaration 2017, https://eosc-portal.eu/sites/default/files/eosc_declaration.pdf (retrieved 2025/12/15)
6. Barend, M., et al. (2017). Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud, Information Services & Use, 37(1):50
7. European legislation on open data, https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data (retrieved 2025/12/15)
8. Commission Implementing Regulation (EU) 2023/138 of 21 December 2022 laying down a list of specific high-value datasets and the arrangements for their publication and re-use (Text with EEA relevance), https://eur-lex.europa.eu/eli/reg_impl/2023/138/oj (retrieved 2025/12/15)
9. FAME project website, https://www.fame-horizon.eu (retrieved 2025/12/15)
10. GDPR (2025, June 5), Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679 (retrieved 2025/12/15)
11. AIA, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM/2021/206 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (retrieved 2025/12/15)

Annex I

Questionnaire completed by each FAME consortium partner

Data Name	
Data description	
Related WP/Task	
Data origin	
Will you re-use any existing data? If yes, how?	
Methodologies for data collection/generation	
Data format	
Where this data will be stored?	
Expected size of the data	
Metadata and Standards used	
For whom might the data be useful?	
Data access, sharing and Licensing	

Annex II

Questionnaire completed by each FAME consortium partner with indexed data assets to the FAME Data Marketplace

Data Asset Name		
Data Asset Overview	**Description**	
	Source/Origin	
	Collection/Generation methodology	
	Type	
	Format	
	Behavior *(Fixed or Changing over time)*	
	Expected size	
Metadata	**Metadata & Standards used**	
	Documentation methodology	
Storage, Backup & Security	**Storage methodology**	
	Backup methodology	
	Security methodology	
Protection & Privacy Provisions	**Privacy & Confidentiality issues**	
	Intellectual property rights responsible	
	Intellectual property rights constraints	
Access & Sharing	**Receiver(s)**	
	Re-use methodology	
	Any other existing data (asset) re-use	
	Access, sharing & licensing methodology	
Archiving & Providing Access	**Preservation planning**	
	Long-term access provision planning	
	Preservation duration	
Responsibility	**Responsible for data asset management**	