D.6.4 Data Management Plan - Final

Document Summary Information

Project Identifier	HORIZON-CL4-2022-DATA-01. Project 101093129		
Project name	Agile and Cognitive Cloud-edge Continuum management		
Acronym	AC^3		
Start Date	January 1, 2023	**End Date**	June 30, 2026
Project URL	www.ac3-project.eu		
Deliverable	D.6.3 Data Management Plan – Final		
Work Package	6		
Contractual due date	30/06/2026	**Actual submission date**	30/06/2026
Type	DMP	**Dissemination Level**	PU
Lead Beneficiary	SPA		
Responsible Author	T. Sarantakos (SPA)		
Contributors	D. Amaxilatis (SPA), N. Tsironis (SPA), T. Sarantakos (SPA), C. Catalán Torrecilla (UCM), A. Rasheed (FIN), L. Garrido (IQU), A. Rehman (IQU), M. Bagaa (FIN)		
Peer reviewer(s)	Th. Tsapikouni (ISI/ATH)		

Funded by the European Union

AC^3 project has received funding from European Union's Horizon Europe research and innovation programme under Grand Agreement No 101093129.

Revision history (including peer reviewing & quality control)

Version	Issue Date	% Complete	Changes	Contributor(s)
v0.1	01/04/2026	10%	Initial Deliverable Structure	D. Amaxilatis (SPA)
v0.2	15/04/2026	25%	Updated Section 2 and 3	N. Tsironis (SPA)
v0.3	29/04/2026	50%	Updated Section 4 and 5	T. Sarantakos (SPA)
v0.4	05/05/2026	70%	Updated Section 7	D. Amaxilatis (SPA)
v0.8	25/05/2026	75%	Internal Review	Th.Tsapikouni (ISI/ATH)
V1.0	15/06/2026	100%	Final Document Editing	D. Amaxilatis (SPA)

Disclaimer

Copyright message

Table of Contents

List of Tables

Glossary of terms and abbreviations used

Abbreviation / Term	Description
AC^3	Agile and Cognitive Cloud-edge Continuum management
AI	Artificial Intelligence
API	Application Programming Interface
ARK	Archival Resource Key
CA	Consortium Agreement
CC 0	Creative Commons Zero licence
CC BY	Creative Commons Attribution licence
CECCM	Cloud Edge Computing and Control Manager
CERIF	Common European Research Information Format
DMLC	Data Management Life Cycle
DMP	Data Management Plan
DDI	Data Documentation Initiative
DOI	Digital Object Identifier
E2E	End to End
EC	European Commission
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Re-usable
FL	Federated Learning
GA	Grant Agreement
GDPR	General Data Protection Regulation
I2C	Inter-Integrated Circuit
IoT	Internet of Things
IPR	Intellectual Property Rights

ML	Machine Learning
NIS Directive	Directive on security of Network and Information Systems
OpenAIRE	Open Access Infrastructure for Research in Europe
PaaS	Platform-as-a-Service
PID	Persistent Identifier
PM	Particulate Matter
SPI	Serial Peripheral Interface
UAV	Unmanned Aerial Vehicle
UC	Use Case
USB	Universal Serial Bus
WP	Work Package
XAI	Explainable Artificial Intelligence

1 Executive Summary

The Final Data Management Plan (DMP) for the AC³ project constitutes the closing milestone in a three-part series initiated at M6 with D6.2 and continued at M18 with D6.3. This document consolidates and finalises the data management methodology adopted throughout the project's three-year lifecycle, reflecting both the decisions established in the earlier versions and the concrete outcomes realised during implementation. The plan maintains its commitment to the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) as the organising framework for all data-related activities. Where the Initial DMP established methodological intentions and the Intermediate DMP elaborated these against an evolving architecture, this Final DMP reports on what was actually built, deposited, and made available to the scientific community and the broader public. Notable developments since D6.3 include the full deployment of the AC³ Data and Service Catalogue, the operational implementation of the Eclipse Data Space Connector (EDC) ecosystem comprising a Cold Data Connector and Streaming IoT Data Connector, and the registration of all three use case datasets within the catalogue under DCAT-AP-compliant semantic descriptions governed by Gaia-X principles. The project's software outputs have been made publicly available through the dedicated GitHub organisation at https://github.com/EU-AC³, and relevant research outputs have been deposited in open repositories in line with the Open Research Data Pilot obligations. Data security and privacy measures described in the previous versions of this document have been consistently applied throughout the project, with particular attention to the video surveillance data of Use Case 2. GDPR compliance has been maintained across all processing activities. The ethical framework has operated as designed, and no substantive deviations from the ethical commitments recorded in earlier DMP versions have been observed. This Final DMP serves simultaneously as a record of what has been achieved and as a reference document for any future exploitation, continuation, or reuse of AC³ data assets beyond the project's formal duration.

2 Introduction

The Data Management Plan is an essential instrument for ensuring the effective, secure, and ethically sound handling of data within a funded research project. For AC³, this document marks the third and final iteration, closing the DMP cycle mandated by the Horizon Europe Guidelines on Data Management. Across its three versions, the DMP has progressively matured from a statement of principles and intentions (D6.2, M6) through an intermediate elaboration informed by the first implementation cycle (D6.3, M18) to the present final account grounded in concrete outcomes (D6.4, M42). The scope of the document spans the full breadth of data produced, collected, and managed within the project: IoT sensor streams from Use Case 1, video and IoT feeds from Use Case 2, astronomical observation data from Use Case 3, and the operational data generated by the AC³ Cloud-Edge Continuum Configuration Management (CECCM) platform itself. It also addresses the outputs made available to the scientific community through publications and software repositories, and the policy and governance structures under which all this data has been handled. The structure of the document deliberately mirrors that of D6.2 and D6.3, preserving continuity and allowing readers familiar with those versions to trace how each section has evolved. Where content is unchanged from a prior version it is retained with minor editorial adjustments; where it has been substantively updated or extended, the new material is clearly integrated. A dedicated subsection (2.3) maps the principal differences from D6.3.

2.1 Mapping AC³ Outputs

The purpose of this section is to map AC³ Grant Agreement (GA) commitments, both within the formal Deliverable and Task description, against the project's respective outputs and work performed.

Table 1: Adherence to AC³ GA Deliverable & Tasks Descriptions

AC³ GA Component Title	AC³ GA Component Outline	Respective Document Chapter(s)	Justification
DELIVERABLE			
DELIVERABLE D6.4 Initial Data management plan			
TASKS			
T6.1 **Data Management**	Implementation and maintenance of the DMP across the project lifecycle	2, 4	The deliverable describes the Data Management Plan of the AC³ project. Therefore, in this deliverable the reader will find: • Open Data and Open Access principles followed

			• Data Management Lifecycle • Data Types collected by project and Use Cases • Data Security and Privacy principles • Ethical Considerations on Data protection
T6.2 **Open Science and Open Access**	Ensuring FAIR compliance and open access to results	3	

2.2 Link with other Project Activities

This deliverable draws on the full body of technical work carried out across all Work Packages. The architectural foundations of the data management framework are defined in D2.2 (Final Release of the CECC framework and CECCM), which establishes the roles of the Catalogue, the Ontology Semantic Reasoner (OSR), and the Life Cycle Management (LCM) component within the overall CECCM architecture. The security and trust framework that governs all data access and exchange within the platform is documented in D2.6 (Security and Trust Management for CECC - Final), which provides the technical basis for the encryption, access-control, and governance mechanisms summarized in Section 5 of this document. The technical implementation of the data management infrastructure (the AC³ Data and Service Catalogue, the EDC-based connectors, and the data processing application add-ons) is described in detail in D3.3 (Initial) and D3.4 (Final) (Report on Data Management for Applications in CECC). These deliverables constitute authoritative technical companions to Sections 4.2.3 and 5.5 of the present document, which address these components from a governance and compliance perspective and are relevant for understanding the evolution of the Catalogue from its first Piveau deployment to its final form. The integration evidence that confirms the data flows described in this DMP have been realised in practice is provided in D5.4 (Final Integration Report), which documents end-to-end validation across all three use cases. The dissemination and exploitation outputs referred to in Section 3.5 are catalogued in full in D6.7 (Dissemination, Communication and Exploitation - Final) and the final update thereof. Together, these deliverables form an integrated body of work within which this DMP serves as the governance and compliance layer.

2.3 Evolution from D6.2 and D6.3

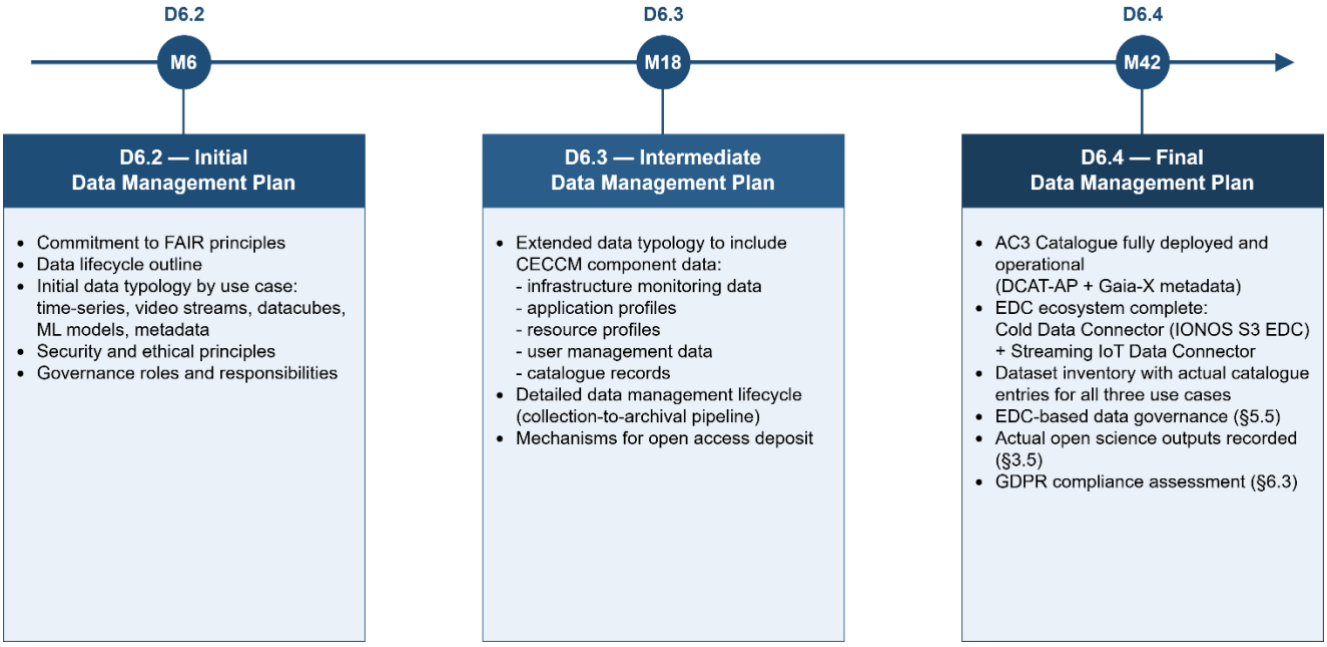

Figure 1 Comparison table showing the key additions at each DMP version: D6.2 (Initial, M6), D6.3 (Intermediate, M18), D6.4 (Final, M42)

The three DMP versions represent a deliberate progression, and understanding the delta between them is important for any reader assessing compliance**. D6.2 established the foundational framework:** a commitment to FAIR principles, an outline of the data lifecycle, an initial typology of data by use case (time-series, video streams, datacubes, ML models, metadata), the security and ethical principles to be followed, and the governance roles and responsibilities. At that stage the architecture was still stabilising and the descriptions of data types were necessarily high-level. **D6.3 built on this foundation in two significant ways:** first, it extended the data typology beyond the use cases to encompass the internal data generated by the CECCM components themselves (infrastructure monitoring data, application profiles, resource profiles, user management data, and catalogue records), reflecting the matured understanding of the platform's data flows; second, it provided a more detailed treatment of the data management lifecycle, including the collection-to-archival pipeline and the mechanisms for open access deposit. **This Final DMP, D6.4,** introduces the following principal updates relative to D6.3. The AC³ Data and Service Catalogue is fully deployed and operational; the dataset inventory described in Section 4.2.3 reflects the actual catalogue entries at project end, including registered datasets for all three use cases with DCAT-AP semantic descriptions under the Gaia-X metadata model. The EDC-based data connector ecosystem is complete: the Cold Data Connector, based on the IONOS S3 EDC extension, is operational for UC3 astronomical data, and the Streaming IoT Data Connector supports UC1. A new subsection (5.5) discusses how these connectors operationalise data governance through enforceable digital contracts. The open science outputs section (3.5) has been added to record what has actually been deposited or published. A GDPR compliance assessment (6.3) documents how the ethical and regulatory commitments stated in prior versions were honoured in practice.

2.4 Deliverable Overview and Report Structure

This deliverable is divided into seven sections:

Section 1 provides the Executive Summary. Section 2 (this section) introduces the document, maps its outputs, and describes its evolution relative to predecessor deliverables. Section 3 covers Open Science and Open Access, including the FAIR data management principles, OpenAIRE compatibility, and the open science outputs actually realised. Section 4 presents the Data Management Plan Overview, including the lifecycle, data typology (use case and CECCM data), and the final catalogue inventory. Section 5 addresses Data Security and Privacy, including security planning, preservation, anonymisation, access policy, and the Gaia-X-based governance framework. Section 6 addresses Ethical Considerations and Data Protection, including a GDPR compliance assessment. Section 7 provides the final Conclusions.

3 Open Science and Open Access

3.1 FAIR Data Principles

The Guidelines on Data Management in Horizon Europe place the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) at the center of data management obligations for funded projects. The approach adopted by AC³, described in detail below, has been consistent across all three DMP versions. This section confirms the implementation of those principles at project end for reuse.

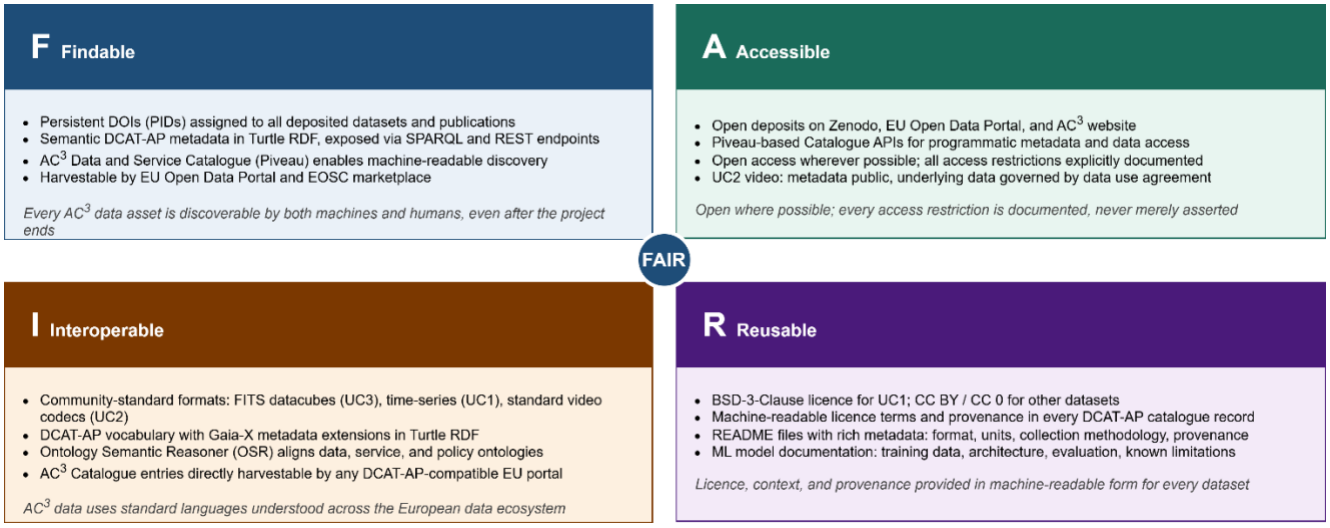

Figure 2 A four-quadrant visual summarising how AC³ implements each FAIR principle

3.1.1 Findable

Any open datasets produced by AC³ are accompanied by documentation that facilitates their understanding and reuse. Persistent Identifiers (PIDs), in the form of Digital Object Identifiers (DOIs), are assigned to deposited outputs. Rich metadata conforming to standard schemes (Dublin Core, CERIF, and DDI where applicable) is maintained for all registered data assets. Datasets follow the naming convention AC³-[WP]-[title]-[VERSION]-[DATE].[TYPE] and carry version numbers that enable users to identify updates through defined change lists. Within the AC³ Catalogue, datasets are further described using semantic DCAT-AP metadata expressed in Turtle RDF, enabling machine-readable discovery across interoperable catalogues. The Piveau-based catalogue infrastructure exposes this metadata through standard SPARQL and REST endpoints, making all registered data assets searchable by external systems and harvestable by portals such as the European Open Data Portal and the EOSC marketplace. This ensures that even data assets that are not publicly downloadable in full are at minimum discoverable through their metadata. The naming conventions, versioning rules, and metadata standards adopted by AC³ have been applied consistently from the project's early stages. This consistency is important for long-term findability: a dataset registered in the AC³ Catalogue today will remain discoverable after the project has ended, provided the Catalogue's hosting infrastructure is maintained, and its DOI will continue to resolve to the deposited record in Zenodo regardless of the project website's longevity.

3.1.2 Accessible

Shared data is deposited in open repositories, accessible through the AC³ website (https://ac3-project.eu) and through external services including Zenodo and the EU Open Data Portal. The AC³ Catalogue itself provides

machine-accessible metadata through its Piveau-based infrastructure, exposing dataset and service descriptions via standard APIs. Access to the underlying data files follows the principle of "as open as possible, as closed as necessary": where no privacy, security, or IPR constraint applies, data is made freely downloadable under an open license. Sensitive data, particularly video recordings from Use Case 2 collected in publicly accessible spaces, is incompatible with full open access due to the privacy constraints imposed by GDPR and the ethical review process. For such data, the metadata record is publicly accessible in the AC³ Catalogue while access to the underlying content is restricted to authorised researchers who have agreed to the applicable data use terms. The justification for each such restriction is documented in the relevant catalogue entry and in this DMP. This approach complies with the Horizon Europe requirement that access exceptions are documented rather than simply asserted. The general principles for handling knowledge and Intellectual Property Rights are governed by the Grant Agreement and the Consortium Agreement, with background and foreground results clearly identified and access rights specified. Background knowledge (pre-existing intellectual property brought into the project by individual partners) remains the property of those partners and is not subject to open access obligations unless otherwise agreed. Foreground results (new knowledge generated within the project) are subject to the open access provisions of the Grant Agreement, and the relevant datasets and publications have been deposited accordingly.

3.1.3 Interoperable

All data produced within the project uses common, standardised, and non-proprietary formats. IoT sensor data from UC1 is available as time-series streams using established IoT data formats. Video data from UC2 follows standard media encoding conventions. Astronomical observations from UC3 are stored and shared as FITS files, the canonical format in the astrophysics community, ensuring interoperability with the full ecosystem of astronomical software tools and data archives. The choice of format in each case has been driven by community standards rather than proprietary or project-specific conventions. The semantic annotations used in the AC³ Catalogue are expressed in Turtle RDF using the DCAT-AP vocabulary and Gaia-X metadata extensions, ensuring compatibility with the European data infrastructure ecosystem. DCAT-AP is the mandatory metadata standard for data portals in the EU, and its adoption means that AC³ catalogue entries can be directly harvested by any DCAT-AP-compatible portal without transformation. The Gaia-X metadata extensions add data sovereignty and federation semantics that go beyond the base DCAT-AP specification, supporting the interoperability requirements of the emerging European data space framework. The Ontology Semantic Reasoner (OSR) within the CECCM further ensures that the data descriptions used by the platform remain aligned with service descriptions, enabling automated reasoning over data and service requirements. The OSR translates high-level policy and intent descriptions into actionable configurations by reasoning over the combined dataset and service ontologies maintained in the AC³ Catalogue. This alignment between data, service, and policy ontologies is a structural guarantee of interoperability within the platform and a precondition for the automated application deployment workflows demonstrated in D5.4.

3.1.4 Reusable

All openly shared datasets carry explicit licensing information. The project has applied the BSD-3-Clause license to the UC1 dataset registered in the AC³ Catalogue, chosen for its permissive terms that allow both academic and commercial reuse with minimal restrictions. Other datasets are governed by Creative Commons licenses (CC BY or CC 0) as appropriate to their nature and the consent basis under which they were collected. README files and rich metadata records provide the contextual information necessary for reuse, including descriptions of data format, measurement units, collection methodology, and provenance. ML models produced within the project are accompanied by documentation sufficient to allow independent validation and reuse. This documentation describes the training data characteristics, the model architecture, the evaluation methodology, and any known limitations or biases. Where models have been made publicly available through the project's GitHub

organisation, they are released under open-source licenses consistent with the data on which they were trained. The AC³ Catalogue's DCAT-AP metadata schema captures licensing, access conditions, and provenance in a structured and machine-readable form for all registered datasets. This means that a consumer accessing a dataset through the Catalogue's API receives not only the data itself but also the machine-readable metadata necessary to assess whether the dataset is suitable for their purposes, under what conditions it may be used, who produced it, and how to cite it. This level of metadata completeness is what distinguishes genuinely reusable data from data that is merely accessible.

3.2 AC³ OpenAIRE Compatibility

AC³ data and publications are made open and offered in compliance with the FAIR data principles and the Open Research Data Pilot, using repositories that are either directly part of OpenAIRE or federatable with it. Zenodo, the primary repository used by AC³ for research data and software deposits, is a core OpenAIRE repository and ensures that all deposited items are automatically indexed in the OpenAIRE portal. This provides AC³ outputs with visibility across the European Open Science Cloud (EOSC) ecosystem without requiring any additional action beyond the initial deposit. Research publications from consortium members have been deposited in open access repositories and published in journals and conference proceedings indexed by OpenAIRE. All deposited publications carry the project's Grant Agreement number (101093129) and the Horizon Europe programme identifier, ensuring that they are correctly attributed to the project in OpenAIRE's funding relationship index. This attribution supports the European Commission's monitoring of open science compliance across Horizon Europe projects and contributes to AC³'s overall compliance record. The project website and the AC³ Catalogue serve as additional discovery points, complementing the OpenAIRE presence. The Catalogue's Piveau infrastructure is built on DCAT-AP, making it technically capable of being harvested by other DCAT-AP-compatible catalogues, including those federated within the EOSC. While direct integration of the AC³ Catalogue with the EOSC marketplace is outside the project's current scope, the adoption of these standards means that such integration could be achieved with minimal additional effort, providing a pathway for the continued discoverability of AC³ data assets after the project's conclusion.

3.3 Publishing Infrastructure for Open Access

Consortium partners have used arXiv and Zenodo as open access publishing and archiving platforms for scientific articles where appropriate, alongside established peer-reviewed journals and conferences in the target domains of cloud-edge continuum computing, data space architectures, AI-driven lifecycle management, and IoT data management. arXiv provides rapid dissemination of preprints with broad visibility in the computing and engineering communities, while Zenodo provides persistent DOI assignment, long-term archiving, and native integration with the Horizon Europe open science requirements. The dual-track open access strategy (green open access through self-archiving within six months of publication acceptance, and gold open access through publication in fully open-access journals or through article processing charges in hybrid journals) has been applied based on each partner's individual circumstances, institutional agreements, and the policies of the target publication venues. Partners with national open access mandates or institutional repository agreements have predominantly followed the green route; others have pursued gold open access where budget allowed. In all cases, the final peer-reviewed manuscript or the published version of record has been made available in an open repository within the Horizon Europe timelines. All training materials, tutorials, and presentations produced within the project have been made publicly available through the AC³ website, SlideShare, and the AC³ YouTube channel, all under Creative Commons Attribution (CC BY) licensing to maximise knowledge reuse and ensure that educational content produced with public funding remains freely accessible to all interested parties. Software repositories on GitHub are released under open-source licenses appropriate to each component, with the license file included in the root of each repository for immediate visibility.

3.4 Publishing Process

The publishing process within the AC³ project follows the principles of open science and open access throughout. At the time of each publication, a machine-readable electronic copy of the final peer-reviewed manuscript or the publisher's version of record is deposited in a trusted repository (primarily Zenodo) under the Creative Commons Attribution International Public License (CC BY) or an equivalent open license. This deposit is performed by the lead author of the publication, with the relevant partners' research managers notified to ensure tracking in the project's dissemination record. Comprehensive metadata for deposited publications is openly accessible, adhering to FAIR principles. The metadata includes author affiliations, publication venue, publication date, the Horizon Europe project identifier and Grant Agreement number, and persistent identifiers both for the publication itself (DOI) and for any associated datasets or software (separate DOIs or software-specific identifiers such as SWHID for Software Heritage). Where publications are accompanied by datasets or software used to produce the results, these are deposited separately with their own identifiers and cross-referenced from the publication's metadata record. Research data accompanying publications has been deposited in Zenodo or equivalent trusted repositories within the timelines specified in the DMP. Open access is provided under CC BY or CC 0 as applicable to the nature of the data. Where certain data could not be made fully open (for instance due to the privacy constraints on video recordings from Use Case 2, or due to ongoing patent applications) the restricted access has been documented in the DMP and in the metadata record of the deposit, with the justification stated clearly. In such cases, the metadata itself remains publicly accessible so that researchers can at least discover the existence of the dataset and contact the data custodian to discuss access conditions.

3.5 Actual Open Science Outputs

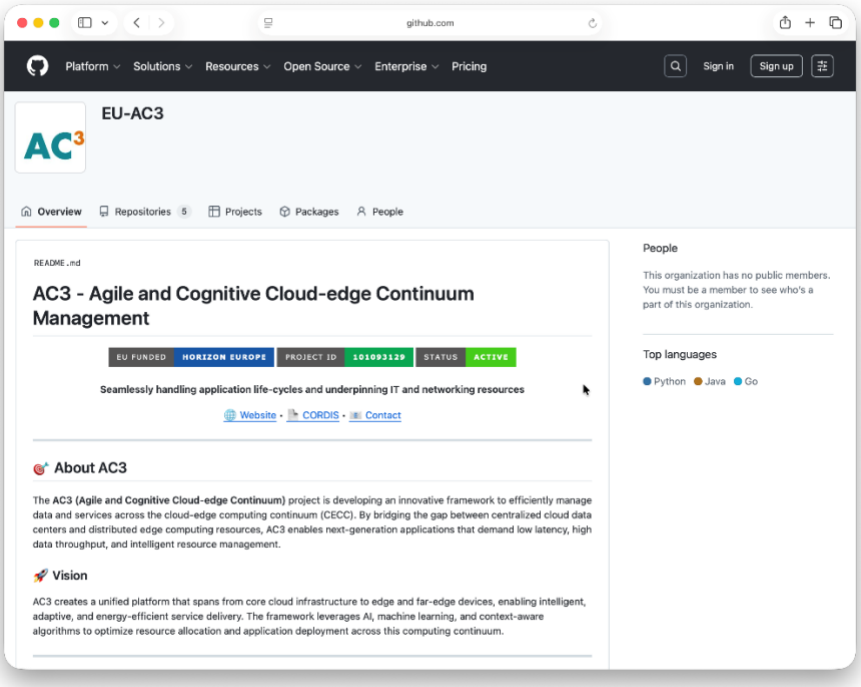

Figure 3 The AC³ GitHub Repository (https://github.com/EU-AC3)

The AC³ consortium has made software components publicly available through the dedicated GitHub organisation at https://github.com/EU-AC3. The repositories published include the EDC extensions developed for the minimum viable AC³ data space, specifically the IONOS S3 EDC extension enabling the Cold Data Connector, as well as supporting components of the CECCM. Each repository is released under an open-source license and includes documentation sufficient for external developers to understand, deploy, and build upon the published components. Research publications from consortium members have been deposited in open access repositories across the project duration, covering topics including the Cognitive Cloud-Edge Continuum framework, AI-driven application lifecycle management, Gaia-X-aligned data space connector architectures, IoT data management at the edge, and astronomical data processing in a cloud-continuum context. A comprehensive list of publications, with titles, venues, DOIs, and open access links, is maintained on the AC³ project website at https://ac3-project.eu and reproduced in D6.6. All three use case datasets have been registered in the AC³ Catalogue with publicly accessible metadata descriptions and, where applicable, deposited in Zenodo with assigned DOIs. The UC3 Cold Data Connector has been validated with the IONOS S3 infrastructure, demonstrating an operational pathway for data consumers to discover, negotiate access to, and retrieve astronomical datasets under policy-governed conditions. This constitutes a proof-of-concept implementation of a Gaia-X-aligned data space that is directly reproducible by other projects working in the European data economy. The full list of publications, datasets, and software outputs, with all persistent identifiers, is maintained in the dissemination tracking system and reported in the final version of D6.6.

4 Data Management Plan Overview

The DMP serves as a comprehensive roadmap for ensuring the efficient and responsible management of research data throughout the lifecycle of our project. This overview provides a high-level summary of our data management practices, highlighting our commitment to promoting data integrity, accessibility, and the long-term usability of valuable research outputs.

4.1 Data Management Life Cycle

The AC³ project's data management lifecycle covers the full journey of data from its generation or collection through to its long-term preservation or disposal. The lifecycle stages are: collection and generation, processing and analysis, storage and curation, sharing and dissemination, and archiving or deletion. These stages are not strictly sequential (in a project like AC³ that involves continuous data streams from live deployments, collection and processing overlap significantly) but they provide a useful organising framework for understanding how data flows through the project and the responsibilities at each stage.

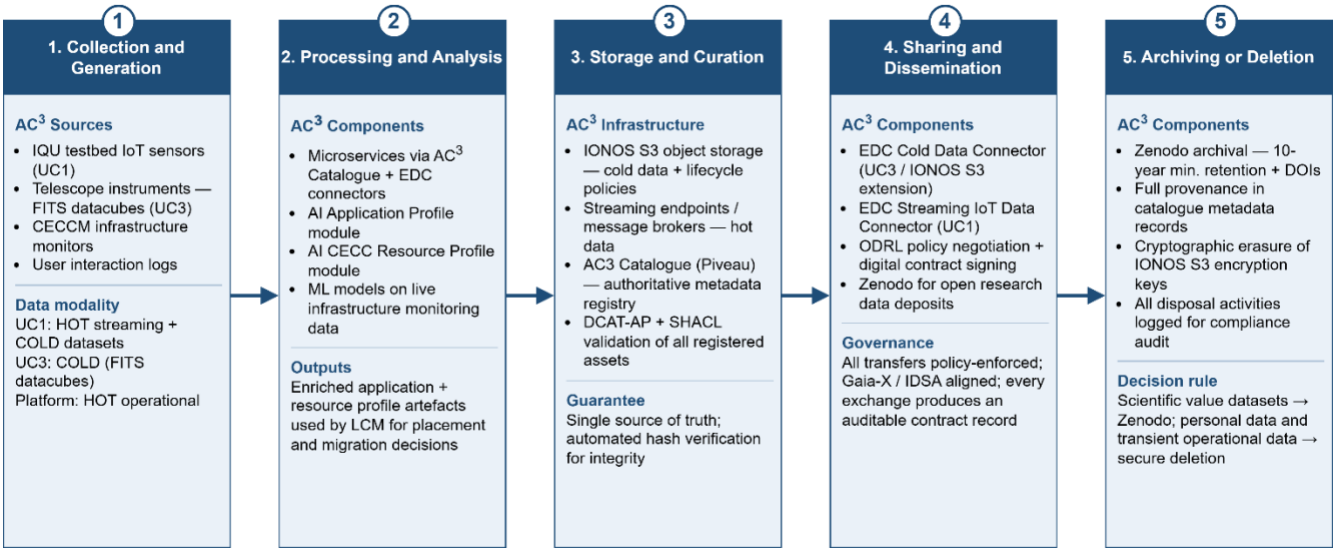

Figure 4 Lifecycle diagram showing the five stages: Collection and Generation, Processing and Analysis, Storage and Curation, Sharing and Dissemination, Archiving or Deletion.

Data collection takes place at multiple levels. At the use case level, UC1 collects IoT sensor readings from the IQU testbed on a continuous basis as hot streaming data, with derived cold datasets constructed from these streams for testing and benchmarking. UC2 collects video feeds from deployed cameras, along with associated IoT device readings, in both hot (live stream) and cold (pre-recorded files) modalities. UC3 handles astronomical observations in the form of FITS datacubes, which are inherently cold data generated from telescope instrumentation. At the platform level, the CECCM continuously generates operational data (infrastructure monitoring metrics, application profiles, resource profiles, and user interaction logs) as part of its normal operation. **Processing and analysis** of the collected data occurs both within the application layer, where deployed microservices consume the data through the AC³ Data and Service Catalogue and the EDC connectors, and within the platform itself, where AI and ML algorithms process monitoring data to drive lifecycle management decisions. The AI-based Application Profile module refines application profiles continuously based on runtime monitoring data; the AI-based CECC Resource Profile module does likewise for infrastructure resource predictions. Both operate on hot operational data and produce enriched profile artefacts that are themselves managed data objects within the platform. **Storage and curation** is handled through a combination of the IONOS cloud infrastructure and the AC³ Catalogue. Hot data is managed through streaming endpoints and

message brokers; cold data is stored in IONOS S3 object storage, with lifecycle policies managing retention and expiry. The Catalogue maintains the authoritative registry of all datasets and their metadata. **Data sharing** is governed by the EDC connector framework, which enforces digital contracts between data providers and consumers before any transfer takes place. **Archiving and deletion** follow documented procedures: valuable datasets are preserved in repositories with long-term availability guarantees, while data that is no longer required is securely deleted with activities logged for audit purposes.

4.2 Data Types and Formats

4.2.1 Purpose of Data Collection and Generation

The following table summarises the data types collected and produced by the three AC³ use cases and the AC³ platform, confirming the typology established in D6.2 and refined in D6.3.

Data Type	Description	UC1	UC2	UC3	AC³ Platform
Time-Series	Measurements or observations recorded over time, allowing analysis of trends, patterns, and anomalies	HOT / COLD	HOT / COLD	N/A	HOT
Video Stream	Live or post-processed continuous flow of audio and video from a deployed camera	N/A	HOT / COLD	N/A	N/A
DataCubes	Three-dimensional array of data values, incorporating two spatial axes representing a region of the celestial sphere	N/A	N/A	COLD	N/A
Machine Learning Models	Files describing mathematical formulas used to make predictions or recognise patterns in data	✓	✓	✓	✓
Metadata	Additional information about collected/generated data, including timestamps, data sources, quality metrics, and contextual information	✓	✓	✓	✓
Operational / Monitoring Data	Infrastructure KPIs, application resource consumption metrics, network link status	N/A	N/A	N/A	HOT

Application / Resource Profiles	AI-generated profiles describing application behaviour and infrastructure characteristics	N/A	N/A	N/A	HOT

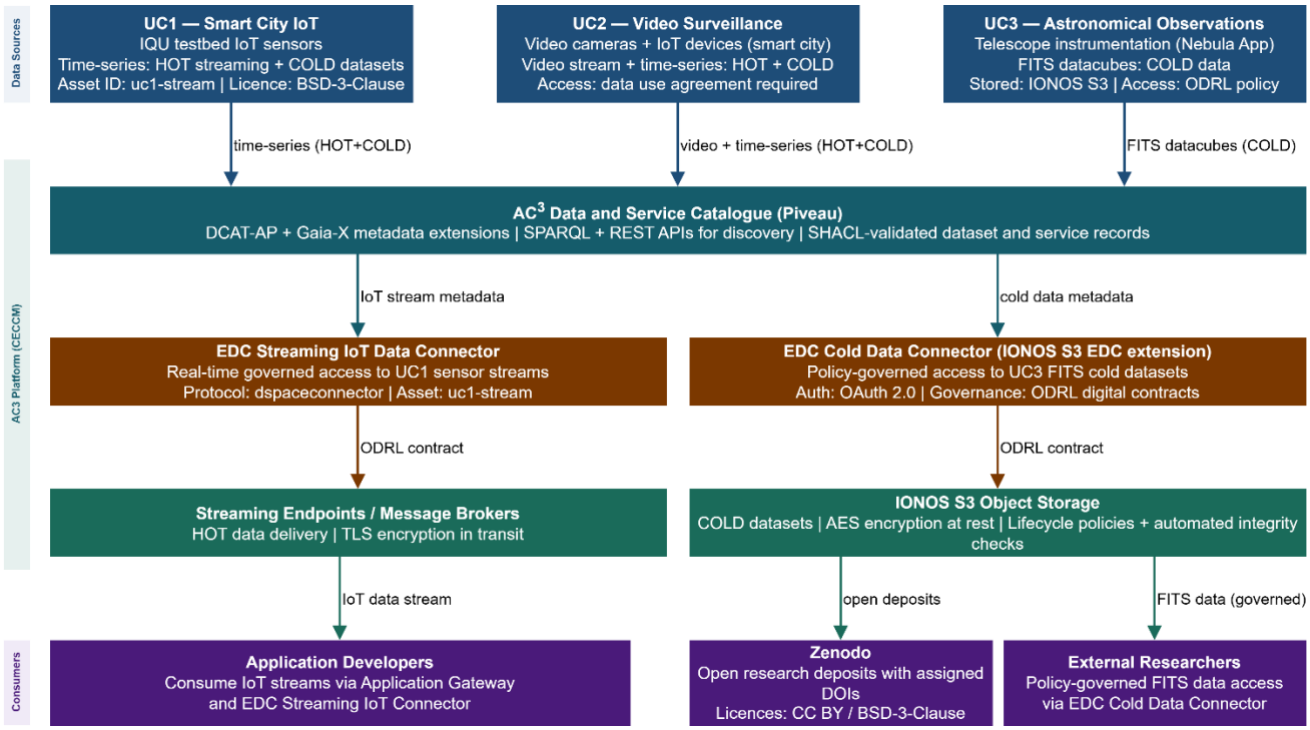

Figure 5 Data flow diagram showing data flowing from the three use cases (UC1, UC2, UC3) down through the CECCM platform layers.

The UC1 dataset registered in the AC³ Catalogue corresponds to sensor data streams originating from the IQU testbed, made available via the Streaming IoT Data Connector under the BSD-3-Clause license. The dataset is assigned the asset ID `uc1-stream` and is accessible via a dedicated data endpoint using the `dspaceconnector` protocol. UC2 produces both live video streams and pre-curated video files suitable for application training; the access conditions for these datasets are governed by the data use agreements established with the relevant data subjects in the deployment environment. The UC3 dataset is registered for the Nebula Astronomy Application, with cold data files in FITS format stored in IONOS S3 and accessible through the Cold Data Connector following successful contract negotiation under the applicable ODRL policy.

4.2.2 CECCM Component Data

The CECCM in its final form manages four categories of internal data, each distinct in nature and lifecycle. Together these categories represent the operational data substrate on which the CECCM's intelligence runs: without accurate, timely infrastructure data, the resource profiling algorithms cannot function; without complete application descriptors and profiles, the LCM cannot make informed placement and migration decisions; without a well-populated catalogue, the OSR cannot reason over data and service dependencies.

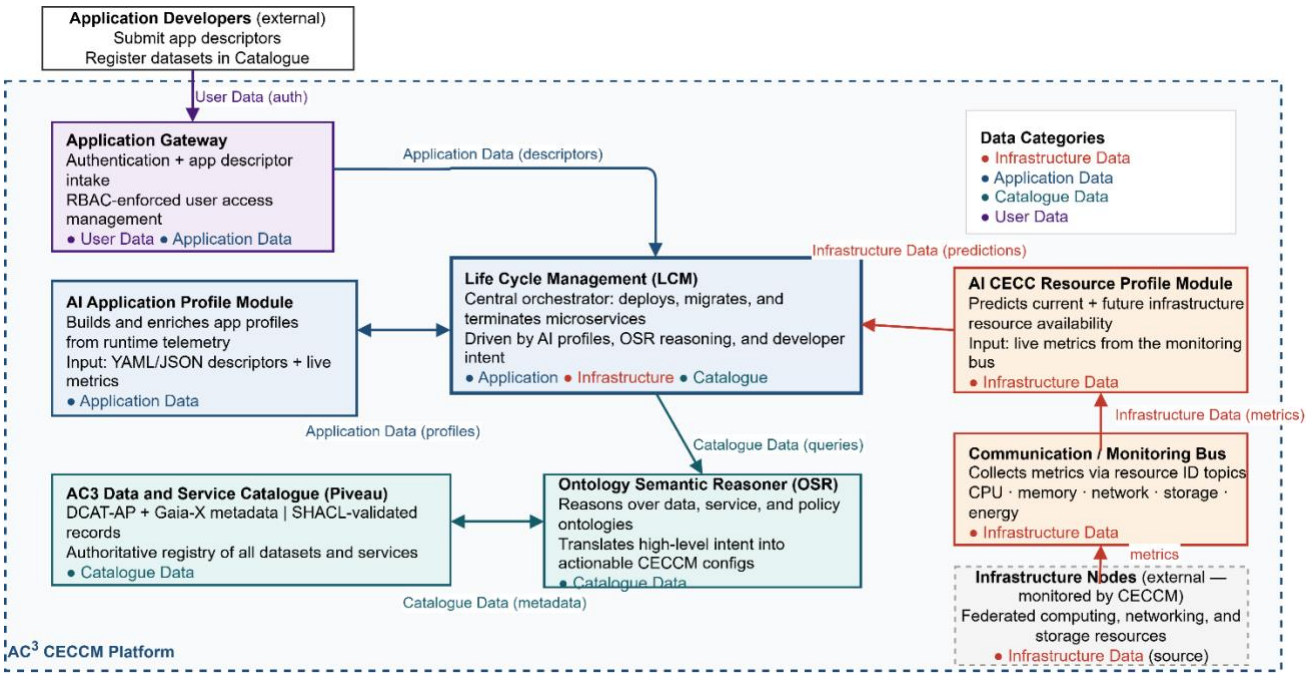

Figure 6 Architectural diagram of the CECCM (adapted from D2.2) with the four data categories (Infrastructure Data, Application Data, Catalogue Data, User Data) showing which CECCM components produce and consume each category.

Infrastructure data encompasses the monitoring metrics collected continuously from the federated computing, networking, and storage resources that make up the CECC. These metrics (CPU utilisation, memory consumption, network bandwidth and latency, storage availability, and energy type indicators) are published over a common communication bus using infrastructure resource identifiers as topics. They feed the AI-based CECC Resource Profile module, which uses ML models to generate both current and predicted resource availability portraits. The monitoring data is inherently short-lived in its raw form; it is aggregated and distilled into profile artefacts that carry a longer lifetime and inform LCM decisions.

Application data covers the descriptors submitted by application developers, the application profiles built and maintained by the AI-based Application Profile module, and the lifecycle management records tracking each microservice's instantiation, migration, and termination. Application profiles are initialised from the developer's application descriptor, which is a structured YAML or JSON document specifying the microservices, their resource requirements, their data dependencies, and their SLA parameters. These profiles are progressively enriched through ML algorithms that monitor runtime behaviour patterns such as traffic flows, data consumption rates, and resource utilisation trends.

Catalogue data is the structured metadata held in the AC³ Catalogue for datasets and services. As described in D3.4, this data is organised under two custom SHACL-defined metadata models (one for datasets and one for services) conforming to DCAT-AP. Each dataset record includes title, description, license, spatial coverage, temporal extent, contact information, asset ID, and references to the service connectors required to access it. Each service record describes the connectors, data mappers, and processing logic that an application needs in order to consume a given dataset, enabling the LCM to automatically assemble a complete application deployment including its data management add-ons.

User data within the CECCM covers the authentication and authorisation records of the application developers who interact with the platform through the Application Gateway. Access is managed through role-based access

control mechanisms, and user records are handled in compliance with GDPR, with data collected to the minimum necessary extent. User data is not shared outside the platform and is retained only for as long as the user maintains an active account.

4.2.3 AC³ Catalogue: Final Dataset Inventory

At project end, the AC³ Catalogue hosts a live and queryable inventory of all datasets and services registered during the project. The Catalogue, deployed on top of the Piveau hub and extended with custom DCAT-AP and Gaia-X metadata models, supports both a web-based human-readable interface and programmatic access via standard SPARQL and REST APIs. The web interface provides advanced search capabilities and allows datasets to be visually linked to the services required to interact with them, facilitating discovery for application developers, data scientists, and external visitors.

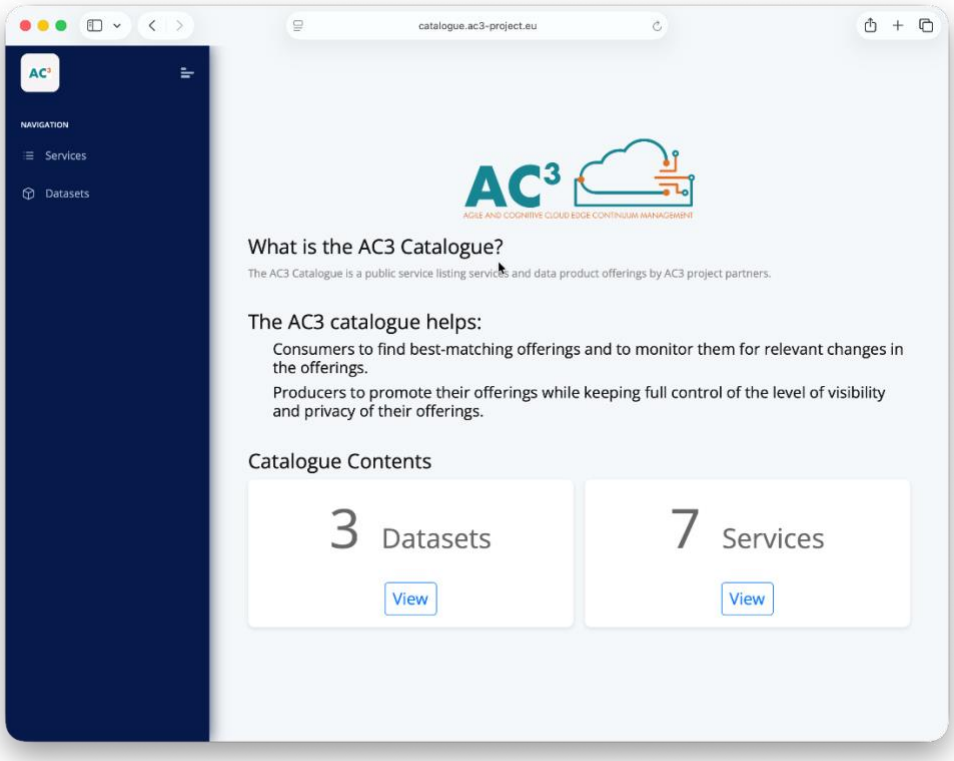

Figure 7 The AC³ Data Space Catalogue (https://catalogue.ac3-project.eu/)

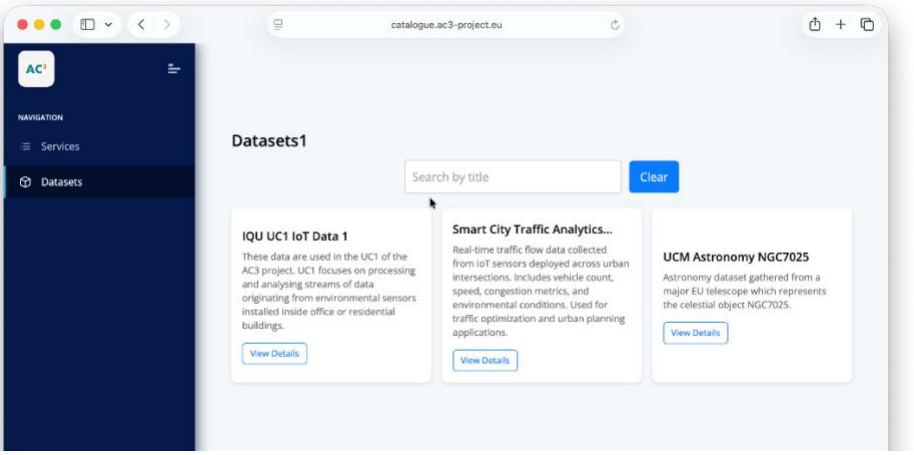

Figure 8 The AC³ Catalogue Dataset List page

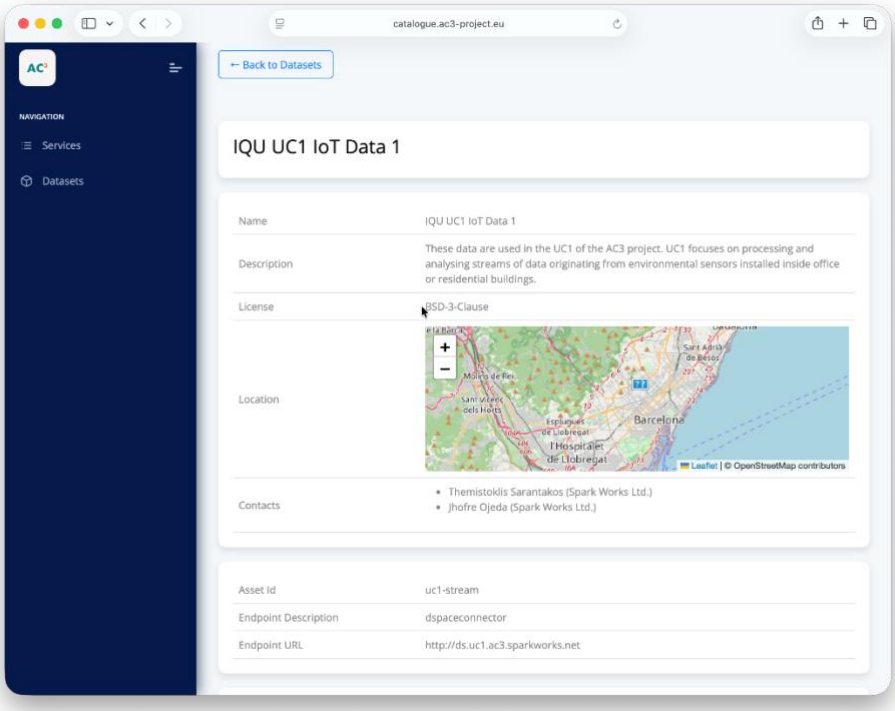

Figure 9 The AC³ Catalogue Dataset View page

Each registered dataset is described with sufficient semantic richness to enable automated discovery and deployment. The dataset schema captures, among other fields, the dataset's title and description, the applicable license, spatial and temporal extent, publisher contact information, the asset ID used within the EDC connector framework, the protocol endpoint for data access, and references to the required connectors and service offerings needed to consume the dataset within the AC³ platform. A corresponding service schema describes each service or connector with equivalent depth. The following code listing shows a representative example of the Turtle RDF description for a dataset registered in the AC³ Catalogue.

```
{
 "@graph": [
  {
   "@id": "https://ac3-project.eu/#Contact1",
   "vcard:hasName": "Jhofre Ojeda",
   "@type": "vcard:Kind",
   "vcard:hasEmail": { "@id": "mailto:j.ojeda@iquadrat.com" },
   "vcard:fn": "Spark Works Ltd."
  },
  {
   "@id": "https://piveau.ac3-project.eu/set/resource/dataset/2befff01-885e-4e5d-a216-e2354baae959",
   "dcat:assetId": "uc1-stream",
   "dcat:contactPoint": [
    { "@id": "https://ac3-project.eu/#Contact2" },
    { "@id": "https://ac3-project.eu/#Contact1" }
   ],
   "https://ac3-project.eu/#requiredServiceOffering": [
    { "@id": "https://piveau.ac3-project.eu/set/resource/service/d8a6b514-3b80-48ec-90cd-d4a229736fb3" },
    { "@id": "https://piveau.ac3-project.eu/set/resource/service/5b6372e6-11a0-47dc-8b1c-761ff81cba25" },
    { "@id": "https://piveau.ac3-project.eu/set/resource/service/115f7472-194d-446c-9518-efb9f55d3f73" }
   ],
   "http://www.w3.org/2003/01/geo/wgs84_pos#lat": { "@value": "41.3976242", "@type": "http://www.w3.org/2001/XMLSchema#decimal" },
   "dct:title": { "@language": "en", "@value": "IQU UC1 IoT Data 1" },
   "dcat:endpointURL": "http://ds.uc1.ac3.sparkworks.net",
   "gx:license": "BSD-3-Clause",
   "schema:name": "IQU UC1 IoT Data 1",
   "https://ac3-project.eu/#requiredConnector": {
    "@id": "https://piveau.ac3-project.eu/set/resource/service/10333e33-0875-4869-9f15-7fc7fccd1d48"
   },
   "dcat:endpointDescription": "dspaceconnector",
   "http://www.w3.org/2003/01/geo/wgs84_pos#long": { "@value": "2.1560832", "@type": "http://www.w3.org/2001/XMLSchema#decimal" },
   "@type": "gx:Dataset",
   "schema:description": "These data are used in the UC1 of the AC3 project. UC1 focuses on processing and analysing streams of data originating from environmental sensors installed inside office or residential buildings."
  },
  {
   "@id": "https://ac3-project.eu/#Contact2",
   "vcard:hasName": "Themistoklis Sarantakos",
   "vcard:fn": "Spark Works Ltd.",
   "@type": "vcard:Kind",
   "vcard:hasEmail": { "@id": "mailto:tsaradakos@sparkworks.net" }
  }
 ],
 "@context": {
  "schema": "https://schema.org/",
  "gx": "https://w3id.org/gaia-x/development#",
  "vcard": "http://www.w3.org/2006/vcard/ns#",
  "dct": "http://purl.org/dc/terms/",
```

```
  "dcat": "http://www.w3.org/ns/dcat#",
  "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 }
}
```

Figure 10 A complete Turtle RDF dataset description from the AC3 Catalogue

This combined Data and Service Catalogue represents a significant evolution from the initial Piveau deployment described in D3.3, and the final implementation as described in D3.4 constitutes the authoritative technical reference. From the DMP perspective, the Catalogue is the primary instrument through which AC³ achieves the findability and accessibility dimensions of FAIR, and through which data governance is operationalised via the connector framework.

4.3 Resources for Data Management

4.3.1 Data Management Roles and Responsibilities

The roles and responsibilities established in D6.2 and carried forward in D6.3 have operated throughout the project as described. They are confirmed here in their final form.

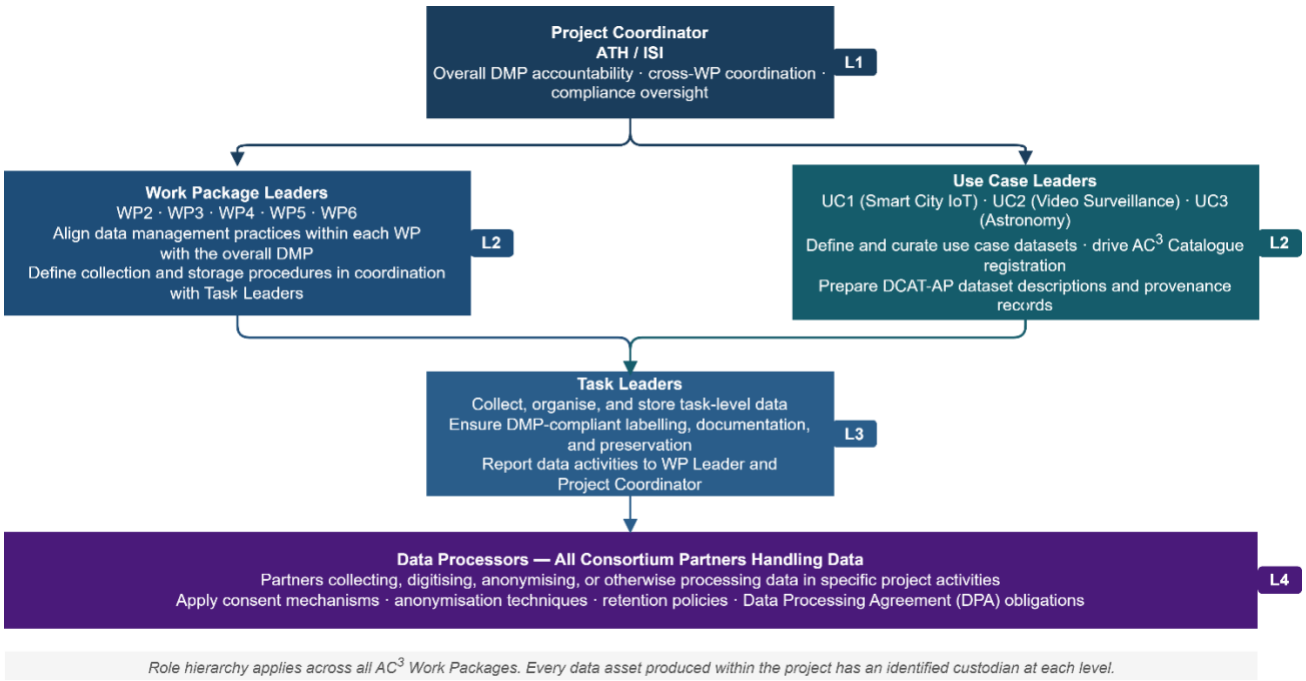

Figure 11 : Data management role hierarchy.

Project Coordinator (ATH/ISI) has overseen and coordinated all data management activities, ensuring compliance with the DMP and serving as the central point of contact for data-related matters across Work Packages. The Coordinator has been responsible for maintaining data integrity at the project level, facilitating cross-WP data sharing, and ensuring that DMP obligations have been fulfilled by the respective partners within the agreed timelines.

Work Package Leaders have held responsibility for aligning data management practices within their respective WPs with the overall DMP. They have worked with Task Leaders and Use Case Leaders to define and monitor data requirements, establish collection and storage procedures, and ensure that data generated within their WP has been properly documented and made available.

Task Leaders have been responsible for collecting, organising, and storing data generated from their assigned tasks, ensuring proper labelling, documentation, and preservation in line with the DMP conventions. They have provided regular updates on data-related activities to their WP Leader and the Coordinator.

Use Case Leaders have driven the definition and curation of the use case datasets, working closely with Task Leaders to ensure that data collected within each use case is accurate, relevant, and aligned with the objectives of that use case. They have played the primary role in preparing dataset descriptions for registration in the AC³ Catalogue.

Data Processors (the project partners who collect, digitise, anonymise, store, or otherwise process data in the context of specific project activities) have been responsible for applying the appropriate consent mechanisms, anonymisation techniques, and retention policies to the data they handle.

5 Data Security and Privacy

Ensuring the security and privacy of data is of paramount importance within the AC³ project. This section outlines the measures and plans in place to safeguard data throughout its lifecycle, confirming the application of the security framework established in D6.2 and D6.3 and documenting the additional governance mechanisms introduced through the EDC-based data exchange infrastructure.

5.1 Data Security Plan

AC³ has implemented a comprehensive data security plan throughout the project. The plan is built on a defence-in-depth philosophy: no single security control is relied upon in isolation, and multiple independent layers of protection are combined so that the failure or circumvention of any one layer does not result in a breach. The plan addresses access control, encryption, network security, and incident response as its four primary pillars.

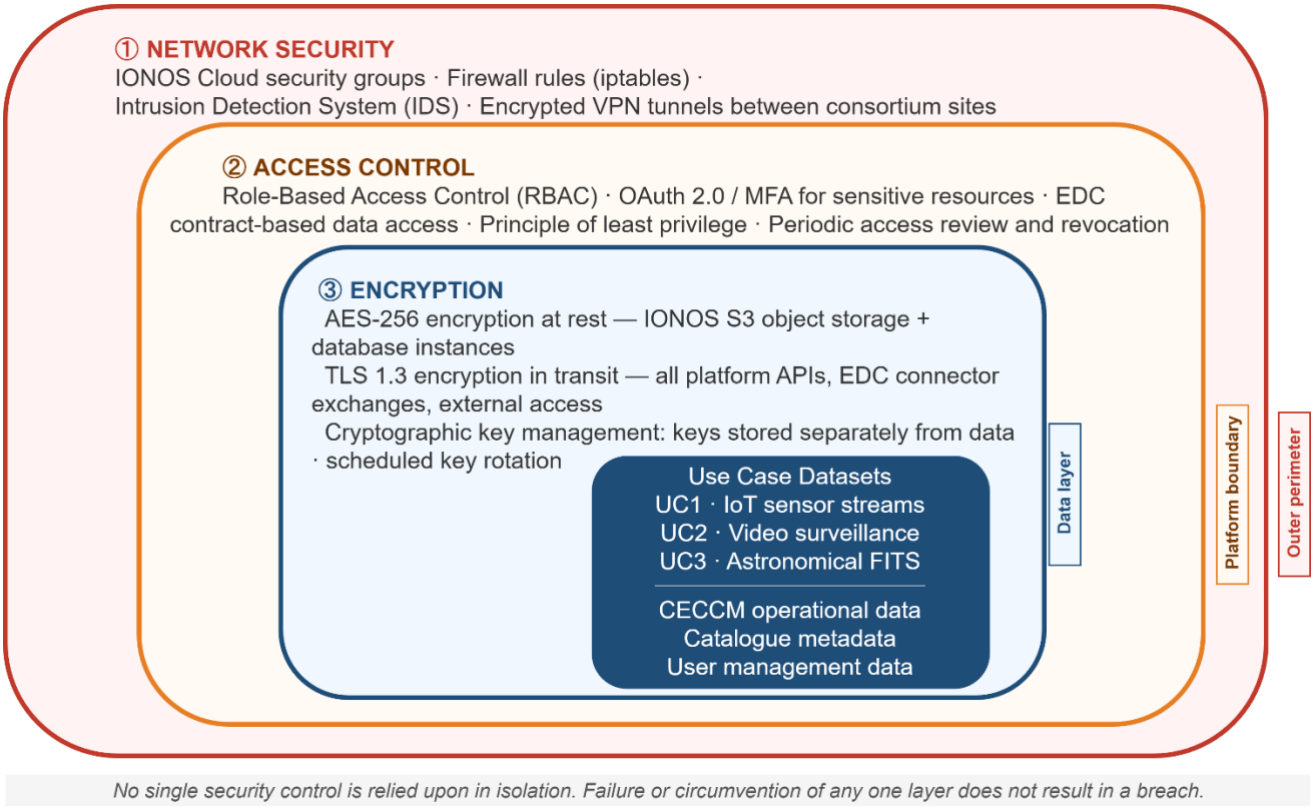

Figure 12 The defence-in-depth security model for AC³ data

Access control is implemented through a combination of user authentication and role-based permissions. Authentication relies on OAuth2 or Multi-Factor Authentication (MFA) depending on the sensitivity of the protected resource and the interface through which it is accessed. Role-based access control (RBAC) is enforced at the database and application levels, ensuring that each user or system component can access only the data and functions necessary for its role. Within the CECCM, this is complemented by the EDC connector framework's contract-based access model, which adds a further layer of policy enforcement at the data-sharing boundary. Encryption is applied to all data at rest and in transit. The IONOS Cloud infrastructure (including IONOS S3 object storage and the secure database services) provides AES-based encryption at rest and TLS-based encryption in transit as native capabilities. Encryption key management follows best practices, with keys stored separately

from the encrypted data and rotated on a defined schedule. The network security perimeter is maintained through firewalls, intrusion detection systems, and regular security assessments. An incident response plan defines the procedures for detecting, reporting, containing, and recovering from security events, including notification of the relevant supervisory authority within 72 hours in the event of a personal data breach as required by Article 33 of GDPR.

5.2 Data Preservation, Archiving and Disposal Plan

The long-term preservation of research data is an obligation that extends beyond the project's formal end date. AC³ has followed a structured plan for data preservation, archiving, and disposal throughout the project, designed to ensure that data of lasting scientific value is available for future reuse while data that is no longer needed is securely removed. Preservation is achieved through a combination of format choice and storage redundancy. Data is preserved in formats that are open, non-proprietary, well-documented, and likely to remain accessible with future technologies: FITS for astronomical data, standard time-series formats for IoT data, and broadly supported video codecs for surveillance recordings. IONOS S3 object storage provides built-in redundancy through data replication across multiple storage nodes. Regular integrity checks (automated hash verification of stored files) provide early detection of data corruption. Backups of the catalogue metadata are maintained separately from the primary storage infrastructure to ensure recovery in the event of a platform failure.

Archiving focuses on preserving datasets of lasting scientific value for future reference and reuse. Use case datasets registered in the AC³ Catalogue are also deposited in Zenodo with assigned DOIs, ensuring their long-term preservation under Zenodo's ten-year minimum retention guarantee. The catalogue metadata records for these datasets include full provenance information (including the project, work package, partners involved, and data collection methodology) enabling future researchers to correctly interpret and cite the archived data. Datasets that are not of lasting scientific value, such as transient operational monitoring data, are not archived but rather subject to the retention policies described below. Disposal of data that is no longer required or legally permissible to retain follows defined procedures. For personal data, the retention period is determined by the purpose for which the data was collected and the applicable legal retention obligations, and deletion is triggered automatically or manually at the end of this period. Secure deletion methods that render data irretrievable (including cryptographic erasure of encryption keys for S3-stored data) are applied in preference to simple file deletion. All disposal activities are logged for audit and compliance purposes, with records retained for a further period to support any subsequent compliance enquiries.

5.3 Data Anonymization

The amount of personal data collected within AC³ is extremely limited given the nature of the project's use cases. Use Cases 1 and 3 involve IoT sensor readings and astronomical observations respectively, and neither category involves personal data in the ordinary sense. The principal area of privacy concern is Use Case 2, which deploys video cameras in smart city environments for crowd surveillance and monitoring. In this context, the video feeds inherently capture images of individuals, making anonymisation a mandatory prerequisite for any use of the data beyond the strictly controlled local processing environment. For all video data from Use Case 2, a systematic anonymisation pipeline has been implemented prior to any storage, sharing, or analysis of the data outside the immediate collection environment. The techniques applied include face detection and blurring using computer vision algorithms trained on publicly available datasets, removal of other potentially identifying biometric features such as gait patterns in contexts where longitudinal tracking might enable re-identification, and where technically feasible, replacement of identifiable individuals with synthetic avatars that preserve the relevant behavioural characteristics for the application's analytical purposes. The effectiveness of the anonymisation pipeline has been validated through a combination of automated re-identification attack testing and manual review. Beyond the video anonymisation pipeline, the project applies the broader principle of privacy by design

to all data collection activities. Data collection instruments (sensors, cameras, software APIs) are configured to collect only the minimum data necessary for the defined purpose, and any fields that could enable individual identification but are not required for the use case are suppressed at the point of collection rather than collected and subsequently discarded. This approach is more robust than post-hoc anonymisation because it eliminates the risk that unanonymised data might be inadvertently retained, copied, or accessed before the anonymisation step is applied.

5.4 Data Access and Use Policy

AC³ implements a robust data access and use policy to govern the appropriate handling, sharing, and utilisation of data. This policy has operated throughout the project and is confirmed in its final form here.

5.4.1 Authorized Access

Only individuals with legitimate research or operational purposes are granted access to data within the AC³ project. Access rights are assigned on a role, responsibility, and need-to-know basis: a partner contributing to Use Case 1 does not automatically have access to the raw video feeds of Use Case 2, and a data processor handling one task's outputs does not have visibility into another task's datasets unless a specific need has been identified and authorised by the relevant Task Leader and Coordinator. The technical mechanisms enforcing this policy are the authentication and authorisation controls described in Section 5.1. At the platform level, the EDC connector framework adds a further dimension of access control: even within the consortium, data transfers through the connector are subject to policy negotiation and contract signing, creating an auditable record of who accessed what data, under what terms, and when. This applies to both internal use case data exchanges and any access by external parties who have been granted conditional access to restricted datasets. Access rights are reviewed periodically and revoked when they are no longer required, for instance when a partner's involvement in a particular task concludes or when a contractor's engagement ends. The principle of least privilege is applied throughout: access is granted to the minimum level necessary for the defined purpose, and elevation to higher access levels requires explicit authorisation from the Coordinator. This approach minimises the attack surface for insider threats and reduces the risk of accidental data exposure through misconfiguration.

5.4.2 Data Sharing

AC³ facilitates data sharing among authorised parties through secure and controlled mechanisms operating at two distinct levels. Within the platform, data sharing between deployed applications and data sources is mediated entirely by the EDC connector framework: no application can access a dataset without first discovering it through the Catalogue, negotiating a contract with the data provider's Connector, and receiving a time-bounded access token. This ensures that every internal data exchange is governed by an explicit, auditable policy agreement. Externally, data is shared through open repositories (primarily Zenodo) under clearly specified licenses and conditions. The license attached to each deposited dataset determines what consumers may do with the data: CC 0 datasets may be used for any purpose without restriction or attribution; CC BY datasets require attribution; BSD-3-Clause software repositories allow redistribution and modification subject to the retention of the copyright notice and the disclaimer. Consumers who require access to restricted datasets (those with limited access due to privacy or IPR constraints) must contact the relevant data custodian and agree to a data use agreement before access is granted. Data sharing agreements, data transfer protocols, and secure communication channels are established for all inter-party exchanges that involve sensitive or restricted data. Within the consortium, these are governed by the Data Processing Agreement (DPA) annexed to the Consortium Agreement, which specifies the roles and obligations of each partner as controller or processor in relation to the personal data they handle. For exchanges with third parties outside the consortium, separate data sharing agreements have been negotiated on a case-by-case basis.

5.4.3 Responsible Use

Users of AC³ data are obligated to adhere to the ethical guidelines and legal requirements governing data use. This obligation is established in the Consortium Agreement for consortium partners, in data use agreements for external researchers accessing restricted datasets, and through the ODRL-expressed policy terms negotiated through the EDC connector for all programmatic data access. The layered nature of these obligations means that responsible use requirements follow the data regardless of the channel through which it is accessed. Responsible use encompasses using data only for the purposes specified at the time of access, maintaining confidentiality for any restricted or sensitive data, refraining from any attempt to re-identify individuals from anonymised datasets, and not repurposing or redistributing data in ways that conflict with the license terms or data use agreement. Violations of these obligations are addressed through the enforcement mechanisms described in Section 5.4.6. All users are informed of these obligations at the point of access, through the catalogue's metadata records, through the terms presented during EDC contract negotiation, and through the data use agreements signed for restricted datasets. Beyond mere legal compliance, responsible data use in the context of AC³ also encompasses a commitment to data quality and integrity. Users who discover errors, inconsistencies, or quality issues in a dataset are expected to report these to the data custodian, supporting the continuous improvement of the data assets maintained in the Catalogue. This collaborative stewardship model recognises that data quality is a shared responsibility between producers and consumers.

5.4.4 Monitoring and Auditing

AC³ employs a multi-layered approach to monitoring data access, use, and compliance. At the infrastructure level, the IONOS Cloud environment provides automated logging of all access to S3 storage and database resources, generating audit trails that record who accessed which data asset, at what time, and from which network origin. These logs are retained for a defined period to support compliance enquiries and security incident investigations. At the platform level, the CECCM's monitoring module provides continuous tracking of data access across all deployed applications. The module records which datasets and data endpoints are accessed by which application instances, the volume and frequency of data transfers, and any anomalous access patterns that might indicate a security event or a violation of the data use policy. These records complement the infrastructure-level logs and provide the application-level context necessary to interpret access events correctly. At the governance level, the EDC connector framework generates a record of every contract negotiation and data transfer, providing an auditable history of all policy-governed data exchanges. These records identify the data provider, the data consumer, the dataset accessed, the policy terms agreed, and the duration of the access period. Together with the infrastructure and platform monitoring data, this creates a comprehensive audit trail that supports both internal data governance oversight and any external compliance review required by the European Commission or national data protection authorities.

5.4.5 Data Governance

AC³ maintains a data governance framework that establishes the organisational structures, processes, and technical controls necessary to ensure the quality, integrity, and appropriate use of data throughout its lifecycle. The governance framework operates at three levels: organisational (roles and responsibilities), processual (data management procedures and lifecycle policies), and technical (the Catalogue, the EDC connectors, and the monitoring infrastructure). At the organisational level, the roles and responsibilities described in Section 4.3.1 define clear accountability for each category of data produced within the project. The Project Coordinator holds overall accountability for DMP compliance; WP Leaders, Task Leaders, and Use Case Leaders hold delegated responsibility for the data generated within their respective domains; and Data Processors are accountable for applying the correct handling procedures to the data they process. This hierarchy ensures that every data asset within the project has an identified custodian responsible for its management. At the technical level, the AC³

Catalogue serves as the authoritative registry of data assets, providing a single source of truth for dataset descriptions, access conditions, and provenance records. The semantic metadata standards applied (DCAT-AP, SHACL shapes, Gaia-X extensions) ensure that governance information is machine-readable and consistently applied, enabling automated compliance checking and reducing the risk of human error in metadata management. The Catalogue's role in the data governance framework extends beyond passive metadata storage: it actively mediates data access through its integration with the EDC connector framework, ensuring that the governance policies expressed in the metadata are technically enforced at the point of data exchange.

5.4.6 Compliance and Enforcement

AC³ enforces its data access and use policy through a combination of preventive technical controls, monitoring, and corrective measures. Preventive controls (authentication, role-based access, encryption, and EDC-governed data contracts) are the primary mechanism, designed to make non-compliant access technically difficult or impossible rather than merely prohibited. Monitoring, as described in Section 5.4.4, provides early detection of potential violations, enabling corrective action before significant harm occurs. Where a violation of the data access and use policy is detected or reported (whether through automated monitoring, internal audit, or external notification), a defined response procedure is followed. The Coordinator is notified, the affected data access is suspended pending investigation, the incident is documented, and appropriate corrective or disciplinary measures are applied. For personal data breaches, the GDPR-mandated notification procedure is triggered: the relevant supervisory authority is notified within 72 hours, and affected data subjects are informed where required. All incidents, even those that do not ultimately meet the GDPR breach notification threshold, are recorded in the project's incident log. Compliance with legal and regulatory requirements (GDPR, the NIS2 Directive on cybersecurity, and any applicable national data protection laws in the countries where project activities take place) is a baseline obligation for all partners. Partners are required to demonstrate compliance as part of their periodic reporting to the Coordinator, and the Coordinator may commission targeted audits of specific partners' data handling practices where concerns arise. The digital contract mechanism of the EDC framework further reinforces compliance by making the acceptance of policy terms an explicit, recorded step in every data exchange, ensuring that no partner or external party can claim ignorance of the applicable conditions.

5.5 EDC-Based Data Governance and Gaia-X Alignment

AC³ has operationalised its data governance through the Eclipse Data Space Connector (EDC) framework, which enables controlled, policy-enforced data transactions between providers and consumers. At the core of this approach is the digital contract: before any data transfer can take place, the data provider's Connector presents a policy expressed in the Open Digital Rights Language (ODRL) specifying the terms under which the data may be used. The consumer's Connector reviews and accepts these terms, after which a contract is signed between the two parties. This contract governs all subsequent data access, including usage rights, data access duration, expiration policies, and revocation mechanisms. While the AC³ platform cannot technically prevent a consumer from misusing data after it has been received, the EDC framework provides mechanisms to detect and terminate the contract in the event of detected violations, ensuring accountability throughout the data exchange lifecycle.

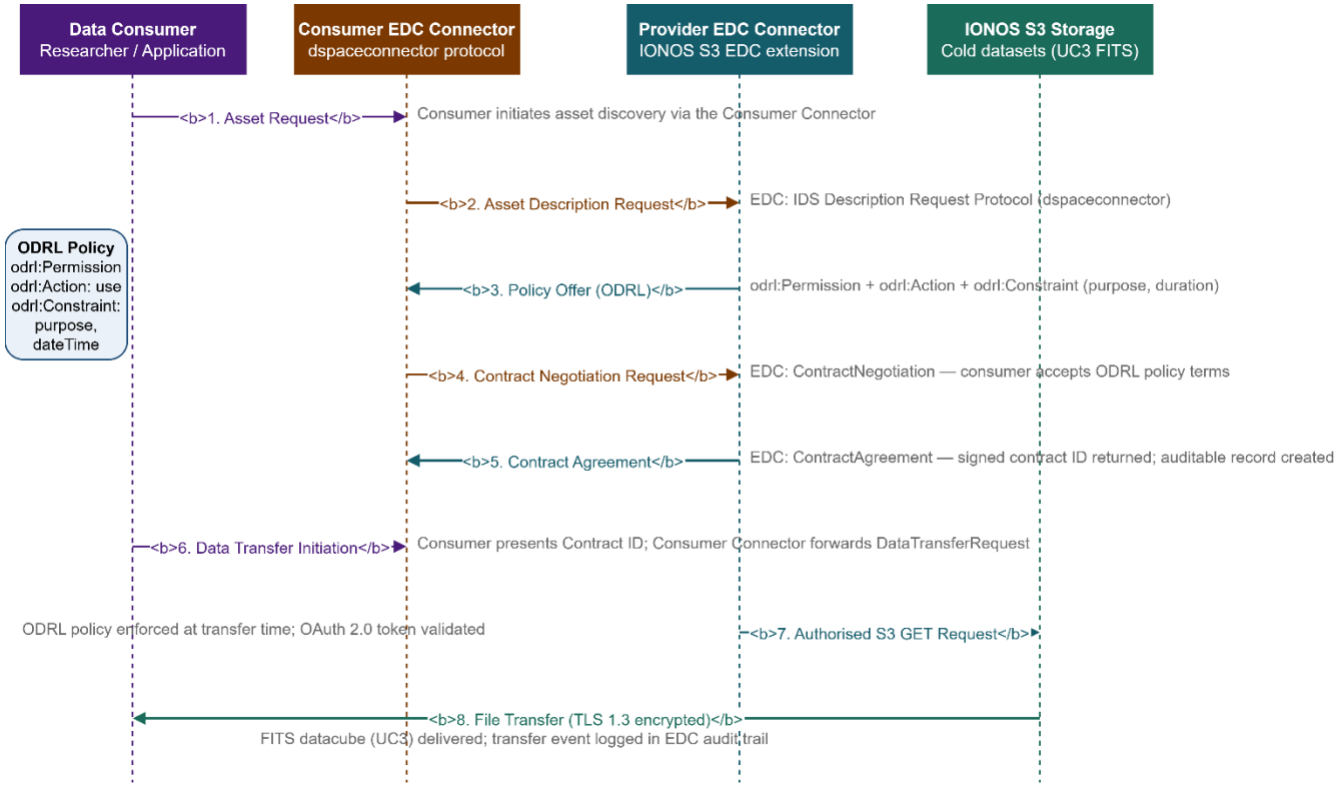

Figure 13 The EDC contract negotiation flow for a cold data access scenario.

```
@prefix odrl:   <http://www.w3.org/ns/odrl/2/> .
@prefix xsd:    <http://www.w3.org/2001/XMLSchema#> .
@prefix dct:    <http://purl.org/dc/terms/> .
@prefix ac3:    <https://ac3-project.eu/policy/> .

# AC3 UC1 Sensor Data Access Policy
# Asset: uc1-stream (temperature, humidity, CO2, IAQ, motion, lux, vibration, battery)
# Provider connector: ionos-s1.sparkworks.net:18182

ac3:uc1-sensor-data-policy
  a            odrl:Set ;
  dct:title    "AC3 UC1 IoT Sensor Data Access Policy" ;
  dct:issued   "2025-01-01"^^xsd:date ;
  dct:valid    "2027-01-31"^^xsd:date ;

  odrl:permission [
    a            odrl:Permission ;
    odrl:target  <https://piveau.ac3-project.eu/set/resource/dataset/2befff01-885e-4e5d-a216-e2354baae959> ;
    odrl:action  odrl:use ;
    odrl:assigner  <https://ionos-s1.sparkworks.net/connector> ;
    odrl:constraint [
      a          odrl:LogicalConstraint ;
      odrl:and   (
```

```
        # Constraint 1 – requesting connector must be an AC3 consortium member
        [ a             odrl:Constraint ;
          odrl:leftOperand  ac3:participantMembership ;
          odrl:operator     odrl:eq ;
          odrl:rightOperand "ac3-consortium" ]

        # Constraint 2 – consumer must declare a permitted research purpose
        [ a             odrl:Constraint ;
          odrl:leftOperand  odrl:purpose ;
          odrl:operator     odrl:isAnyOf ;
          odrl:rightOperand ac3:AnomalyDetection,
                            ac3:BuildingEnergyOptimization,
                            ac3:ResearchAndDevelopment ]

        # Constraint 3 – access expires at end of project lifetime
        [ a             odrl:Constraint ;
          odrl:leftOperand  odrl:dateTime ;
          odrl:operator     odrl:lteq ;
          odrl:rightOperand "2027-01-31T23:59:59Z"^^xsd:dateTime ]
      )
    ]
] ;

odrl:prohibition [
    a           odrl:Prohibition ;
    odrl:target   <https://piveau.ac3-project.eu/set/resource/dataset/2befff01-885e-4e5d-a216-e2354baae959> ;
    odrl:action   odrl:commercialize
] ;

odrl:obligation [
    a               odrl:Duty ;
    odrl:target       <https://piveau.ac3-project.eu/set/resource/dataset/2befff01-885e-4e5d-a216-e2354baae959> ;
    odrl:action       odrl:attribute ;
    odrl:attributedParty <https://ac3-project.eu>
] .
```

Figure 14 A representative ODRL policy expression used in the AC³ EDC connector configuration.

The AC³ Cold Data Connector, based on the IONOS S3 EDC extension, realises this governance framework for stored cold datasets, including the UC3 astronomical data registered in the Catalogue for the Nebula Astronomy Application. It handles metadata registration in the AC³ Catalogue, contract negotiation using OAuth 2.0 for authentication, and data transfer from IONOS S3 object storage to the consumer via the agreed endpoint. The Streaming IoT Data Connector applies equivalent governance to the real-time data streams of UC1, mediated through the `dspaceconnector` protocol endpoint registered in the Catalogue under the asset ID `uc1-stream`. The overall data management architecture aligns with the Gaia-X federated model, under which data sovereignty remains with the provider and access is granted on the basis of well-defined, enforceable agreements. Compliance with GDPR is embedded at multiple levels (storage, transfer, and access management) ensuring that personal data is handled in accordance with the principles of lawfulness, fairness, and transparency. The Piveau-based AC³ Catalogue enhances this framework by providing structured, semantic metadata descriptions that

make licensing terms, privacy restrictions, and compliance indicators visible and machine-readable to all users of the catalogue. This governance model also aligns closely with the International Data Spaces Association (IDSA) Reference Architecture Model: AC³'s data sovereignty approach (fine-grained access control, usage policy enforcement, data lifecycle management, and auditable logging) maps directly to the foundational principles of the IDSA framework, positioning the project's infrastructure as a reference implementation of European data space principles in the Cognitive Cloud-Edge Continuum domain.

6 Ethical Considerations and Data Protection

Ethical considerations and data protection are fundamental aspects of the AC³ project. This section addresses the ethical dimensions of data management, privacy concerns, and compliance with the General Data Protection Regulation, and introduces a compliance assessment reflecting the project's performance against its ethical commitments.

6.1 Ethical Aspects of Data Management

AC³ upholds ethical principles in all stages of data management to ensure the rights and interests of individuals and organisations are respected. The ethical framework adopted by the project is grounded in the four principles set out in the European Code of Conduct for Research Integrity (reliability, honesty, respect, and accountability) and applied specifically to the data management context through the provisions described in the following subsections.

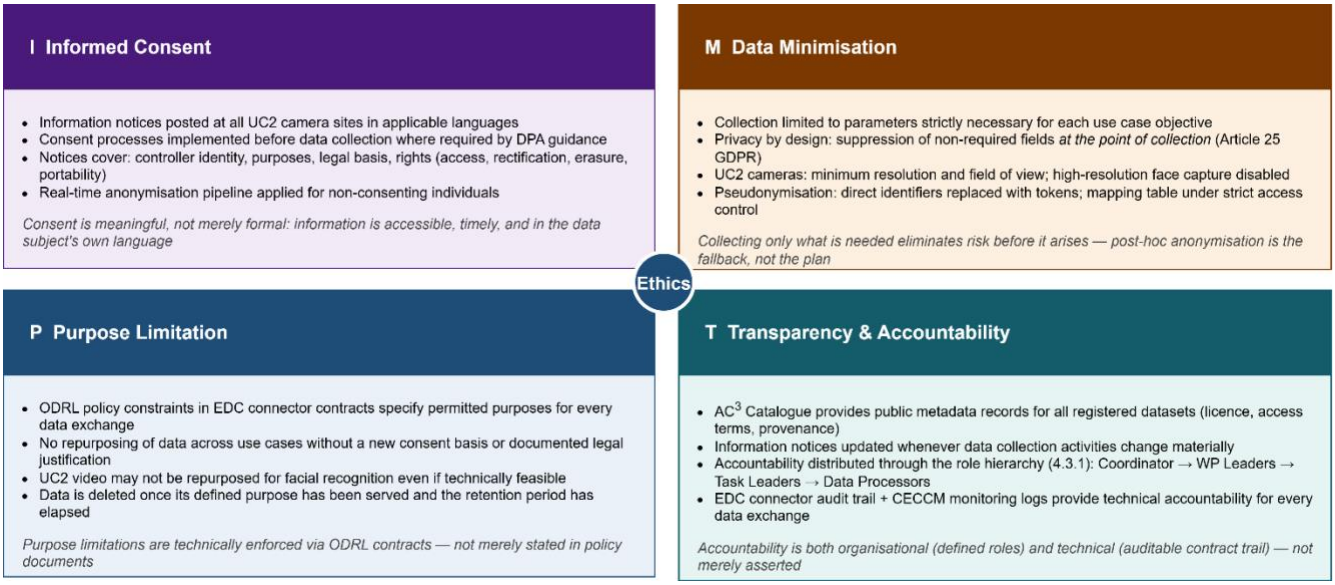

Figure 15 The four core ethical principles to the specific AC³ mechanisms that implement them. Informed Consent, Data Minimisation, Purpose Limitation and Transparency & Accountability.

6.1.1 Informed Consent

AC³ ensures that individuals whose data is collected have the necessary information and understanding to provide informed consent for those data management activities. This obligation is most directly relevant to Use Case 2, where video cameras deployed in smart city environments capture images of individuals who may not be directly party to the research activity. In this context, appropriate information notices have been posted in the areas covered by the cameras, and where direct consent was required by the applicable data protection authority guidance, consent processes were implemented before data collection commenced. The content of the information provided to potential data subjects covers the identity and contact details of the data controller, the purposes of the data collection, the legal basis on which it is conducted, the categories of data collected, the retention period, the recipients of the data, and the rights of the data subject under GDPR. These rights include the right to access, rectification, erasure, restriction of processing, data portability, and the right to object. This information has been made available in the official languages of the countries where the Use Case 2 pilots operate, ensuring genuine accessibility rather than formal compliance. Where individuals have objected to the collection of their image data, or where the applicable legal framework does not provide a sufficient basis for

processing without explicit consent in the specific deployment context, technical measures (specifically the real-time anonymisation pipeline described in Section 5.3) have been applied to ensure that the data of non-consenting individuals is anonymised before any retention or further processing. This approach preserves the scientific value of the dataset while fully respecting the rights of individuals who have not chosen to participate.

6.1.2 Data Minimization

AC³ follows the principle of data minimisation throughout all data collection activities, limiting the collection of personal data to what is strictly necessary for the defined research purposes of each task and use case. This principle has been operationalised at the design stage, through the selection and configuration of data collection instruments, rather than applied as a post-hoc filter, reflecting the privacy by design approach mandated by Article 25 of GDPR. In practice, data minimisation means that sensor configurations in Use Cases 1 and 3 are set to collect only the measurement parameters required by the relevant application scenarios, suppressing any additional telemetry that the hardware is capable of generating but that is not needed for the research objectives. For Use Case 2, camera configurations are set to the minimum resolution and field of view needed for the surveillance and monitoring purposes of the use case, and any features capable of capturing audio or identifying specific individuals beyond what is necessary for crowd density estimation (such as high-resolution face capture) are disabled or anonymised at the point of collection. Advanced techniques such as data anonymisation and pseudonymisation are employed to further protect individual identities wherever they cannot be fully excluded from collection. Pseudonymisation replaces direct identifiers with non-identifying tokens in the stored dataset, with the mapping table maintained separately under strict access controls. This allows the data to be used for research purposes without the risk that the stored dataset, if inadvertently disclosed, would directly identify individuals. Anonymisation (where technically achievable with a sufficient level of assurance) goes further by making re-identification computationally infeasible even with access to auxiliary data sources. The data minimisation approach has been reviewed at each project milestone to ensure that it remains appropriate as the data collection activities evolve.

6.1.3 Purpose Limitation

Data collected within AC³ is used solely for the purposes established at the time of collection and communicated to data subjects through the consent and information mechanisms described in Section 6.1.1. This principle is technically enforced through the ODRL-based policy terms embedded in the EDC connector contracts: each data contract specifies the permitted purposes of use, and the contract management system maintains a record of the agreed purposes that can be audited at any time. No repurposing of personal data is permitted without either a new consent basis or a legal justification that is documented and reviewed by the Coordinator. Purpose limitation operates in two dimensions within the AC³ project. First, it constrains how data collected for one use case may be used in another: Use Case 2 video data collected for smart city crowd monitoring may not be repurposed for facial recognition research, even if such a repurposing might be technically feasible with the available data. Second, it constrains how data collected for research purposes within the project may be used commercially after the project ends: any commercial exploitation of the datasets must be grounded in a legitimate basis (typically a new consent or a legitimate interest assessment) and must be disclosed to data subjects where required. The purpose limitation principle also shapes the data lifecycle more broadly. Once a dataset has served the purpose for which it was collected and the applicable retention period has elapsed, it is deleted rather than retained for potential future use. This approach avoids the accumulation of data holdings whose future purposes are speculative, reducing both the privacy risk and the administrative burden of maintaining data that no longer serves a defined function. The exceptions to this general principle (datasets of lasting scientific value that are archived for future research use) are governed by the archival provisions of Section 5.2 and are subject to appropriate governance measures that ensure their continued responsible management.

6.1.4 Transparency and Accountability

AC³ promotes transparency in its data management practices through multiple channels operating at different levels of the stakeholder hierarchy. At the level of research data governance, the AC³ Catalogue provides a publicly accessible record of all datasets registered in the project, including their metadata, licensing conditions, and access terms. This means that any member of the public can discover what data the project has collected and produced, understand under what conditions it is available, and identify who to contact with questions or concerns. At the level of individual data subjects (particularly those whose images are captured by the UC2 cameras) transparency is provided through the information notices and consent mechanisms described in Section 6.1.1. These notices and mechanisms are reviewed and updated whenever the data collection activities change in a way that materially affects the information that was originally provided. Where changes are significant, data subjects who provided consent under the original terms are informed and given the opportunity to withdraw their consent or provide a fresh consent to the revised terms. Accountability for data management practices rests with the Project Coordinator and is distributed through the role hierarchy described in Section 4.3.1. The Coordinator is responsible for maintaining the DMP and ensuring that all partners adhere to it; WP Leaders, Task Leaders, and Data Processors are accountable within their respective domains. The governance structures described in Section 4.3.1 and the technical enforcement mechanisms described in Section 5.5 together provide both organisational and technical accountability. The former operates through defined roles and periodic audits; the latter through the EDC connector's auditable contract trail and the CECCM's monitoring logs. This dual accountability structure ensures that responsible data management is not merely a matter of stated policy but is embedded in the operational practices and technical infrastructure of the project.

6.2 Data Protection and Privacy

AC³ prioritises data protection and privacy to safeguard the confidentiality, integrity, and availability of data throughout its lifecycle. The data protection approach adopted by the project translates the high-level ethical commitments described in Section 6.1 into concrete technical and organisational measures, consistent with the requirements of GDPR and the NIS2 Directive on cybersecurity. Secure data storage is implemented through the IONOS Cloud infrastructure, which provides AES encryption at rest for all stored data, including the S3 object storage used for cold datasets and the database instances used for operational and catalogue data. Access controls and user authentication mechanisms ensure that only authorised individuals and system components can access the stored data. The separation of encryption key management from data storage (a fundamental principle of cryptographic data protection) is maintained throughout, so that access to the physical storage medium does not imply access to the plaintext data. Data transfer and sharing is protected through TLS encryption for all data in transit, applied consistently across internal platform communications, EDC connector exchanges, and external access to the Catalogue's APIs. Where data is transferred to external researchers or deposited in public repositories, the channel of transfer is either a secure API connection (for Zenodo deposits) or an encrypted file transfer, and the data is accompanied by the metadata and license information necessary for its correct interpretation and use. The EDC connector's contract-based access model provides an additional layer of transfer-time governance, ensuring that the policy terms agreed during contract negotiation remain in force throughout the data exchange process. Compliance with GDPR is embedded throughout the project's data management practices. All data processing activities that involve personal data have been assessed against the GDPR's lawfulness requirements, and the applicable legal basis for each processing activity has been documented. Individual rights (access, rectification, erasure, restriction, portability, and the right to object) have been operationalised through defined procedures available to any data subject who wishes to exercise them. Privacy Impact Assessments have been conducted for processing activities with a high risk to individuals' rights and freedoms, and the mitigation measures identified in those assessments have been implemented. The full compliance assessment is provided in Section 6.3.

6.3 GDPR

GDPR Requirement	GDPR Article	Status	Evidence / Implementation Notes
Lawful Basis for Processing	Art. 6	☑ COMPLIANT	Record of Processing Activities (RoPA) maintained by Coordinator; informed consent for UC2 pilot participants; legal necessity for UC1 and UC3 research activities; legal basis documented for each processing activity
Data Minimisation	Art. 5(1)(c)	☑ COMPLIANT	Privacy by design applied at collection stage (Article 25 GDPR); non-required sensor fields suppressed at source; UC2 cameras set to minimum necessary resolution; pseudonymisation applied where full anonymisation is not feasible
Purpose Limitation	Art. 5(1)(b)	☑ COMPLIANT	ODRL policy constraints in EDC connector contracts specify and enforce permitted purposes for every data exchange; no cross-use-case repurposing permitted without new consent basis; purpose is technically enforced, not merely stated
Data Subject Rights	Arts. 15–22	☑ COMPLIANT	Defined procedures for: access, rectification, erasure, restriction of processing, data portability, and right to object; communicated to data subjects via information notices in applicable languages; CECCM supports data deletion on request
Technical Security Measures	Art. 32	☑ COMPLIANT	AES-256 encryption at rest (IONOS S3); TLS 1.3 in transit for all APIs and EDC transfers; RBAC + OAuth 2.0 / MFA; IONOS Cloud security groups and firewall rules; IDS; keys managed separately from data with rotation schedule
Privacy Impact Assessment (PIA)	Art. 35	☑ COMPLIANT	PIA conducted for UC2 video surveillance (highest-risk processing activity); four-step methodology: describe processing → assess necessity/proportionality → identify risks → mitigate; all identified mitigations implemented (anonymisation pipeline, access controls, retention policies)
Data Processing Agreement (DPA)	Art. 28	☑ COMPLIANT	DPA annexed to the Consortium Agreement; controller and processor roles defined for all partners; obligations for data handling, confidentiality, sub-processing, and audit rights specified; all processing within EEA (IONOS and Arsys infrastructure)
Breach Notification Procedure	Art. 33	☑ COMPLIANT	72-hour supervisory authority notification procedure defined and tested; incident log maintained; data subjects notified where required; no data breaches reported during the project — attributed to defence-in-depth security posture and EDC-based governance

Overall Assessment: All 8 GDPR requirements COMPLIANT
No personal data breaches reported during the AC3 project duration (2022–2026)
Processing exclusively within the EU

Figure 16 AC^3's performance against the key GDPR requirements: Lawful Basis, Data Minimisation, Purpose Limitation, Data Subject Rights, Security Measures, PIA Conducted, DPA in place, and Breach Notification Procedure.

All processing of personal data within the project has been conducted on one of the lawful bases identified at the project's outset: informed consent for data subjects participating in the UC2 pilot environment, necessity for the performance of the research activity as defined in the Grant Agreement, or compliance with a legal obligation. No personal data has been stored in a form that permits identification of individual users beyond what is strictly necessary for the purposes of the relevant activity. The mapping of each processing activity to its legal basis has been maintained in a Record of Processing Activities (RoPA) document held by the Coordinator and updated whenever a new processing activity is introduced. Privacy Impact Assessments were conducted for the data processing activities with the greatest potential to affect individuals' rights and freedoms, in particular the video surveillance component of Use Case 2. The PIAs followed the four-step methodology recommended by the relevant data protection authorities: describing the processing, assessing necessity and proportionality, identifying and assessing risks to data subjects, and identifying measures to mitigate those risks. The technical measures arising from these assessments (the anonymisation pipeline described in Section 5.3, strict purpose-limited access controls, automated retention and deletion policies, and the information notice and consent mechanisms described in Section 6.1.1) have been implemented and maintained throughout the project. International data transfers within the consortium (involving partners in multiple EU Member States) have been compliant with GDPR requirements, with processing taking place exclusively within the European Economic Area. The IONOS and Arsys cloud infrastructure used by the project is located within the EEA, ensuring that no personal data is transferred to third countries without the safeguards required by Chapter V of GDPR. All partners have maintained their obligations as data controllers or data processors, as defined in the Data Processing Agreement annexed to the Consortium Agreement, and have documented their processing activities in accordance with

GDPR accountability requirements. No data breaches have been reported during the project duration. The incident response plan described in D6.2 has remained available and tested but has not been required for activation, an outcome attributed to the defence-in-depth security posture and the EDC-based governance framework maintained throughout the project.

6.3.1 Lawful Basis for Data Processing

AC³ identifies and applies the appropriate lawful basis for data processing, such as consent, legitimate interest, or contractual necessity.

6.3.2 Individual Rights

AC³ respects the rights of individuals under the GDPR, including the right to access, rectify, erase, and restrict the processing of their personal data.

6.3.3 Data Portability

AC³ facilitates data portability, allowing individuals to obtain and transfer their personal data in a structured, commonly used, and machine-readable format upon request.

6.3.4 International Data Transfers

AC³ ensures that any transfer of personal data to countries outside the European Economic Area (EEA) is done in compliance with GDPR provisions, such as using appropriate safeguards like standard contractual clauses or obtaining individual consent.

6.3.5 Accountability

AC³ maintains documentation, records, and policies demonstrating compliance with the GDPR. Regular audits and internal reviews are conducted to assess and improve data protection practices.

By incorporating ethical considerations and adhering to data protection regulations like the GDPR, AC³ strives to ensure responsible and trustworthy data management practices while safeguarding individuals' privacy rights.

7 Conclusions

The Final Data Management Plan for AC³ closes the three-version DMP cycle mandated by Horizon Europe. Across D6.2, D6.3, and this document, the project has demonstrated a consistent and progressively deepened commitment to responsible, FAIR-compliant, and ethically sound data management. The foundational principles established at the outset (FAIR compliance, open access to the extent possible, GDPR adherence, data minimisation, and clear governance roles) have been applied throughout the project's three-year duration. What has grown across the three versions is the technical concreteness of these principles: from intentions stated in D6.2 to architecture described in D6.3 to implementation evidenced in D6.4. The AC³ Data and Service Catalogue, the EDC-based connector ecosystem, the Gaia-X-aligned governance model, and the datasets registered and made discoverable through these instruments are the tangible outcomes of the data management methodology. All three use case datasets have been registered in the AC³ Catalogue with rich semantic metadata. Software components have been made publicly available through the project's GitHub organisation. Research outputs have been deposited in open repositories. The CECCM's internal data flows have operated under the security, privacy, and retention policies set out in this plan. No GDPR breaches have occurred, and the ethical commitments made in earlier versions of the DMP have been honoured in practice. The AC³ project's DMP has served its intended purpose: not merely as a compliance document, but as an active instrument guiding how data was designed, collected, managed, governed, and ultimately shared with the wider scientific and technical community. The data assets and infrastructure created within the project are available for future exploitation, and the governance model demonstrated provides a replicable reference for data space implementations within the European cognitive cloud-edge continuum landscape.

8 References

[1] "AC³ Website," [Online]. Available: https://ac3-project.eu.

[2] European Parliament and the Council of Europe, Regulation (EU) 2016/679 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, 2016.

[3] European Parliament and the Council of Europe, Directive 2022/2555 on measures for a high common level of cybersecurity across the Union, 2022.

[4] "Open Research Europe," [Online]. Available: https://open-research-europe.ec.europa.eu/.

[5] AC³ Consortium, "D6.2 Data Management Plan - Initial," Deliverable, 2023.

[6] AC³ Consortium, "D6.3 Data Management Plan - Intermediate," Deliverable, 2023.

[7] AC³ Consortium, "D3.4 Report on Data Management for Applications in CECC - Final," Deliverable, 2025.

[8] AC³ Consortium, "D2.2 Final Release of the CECC Framework and CECCM," Deliverable, 2023.

[9] AC³ Consortium, "D2.6 Security and Trust Management for CECC - Final," Deliverable, 2023.

[10] Eclipse Foundation, "Eclipse Dataspace Connector," [Online]. Available: https://github.com/eclipse-edc.

[11] "Piveau Data Management Infrastructure," [Online]. Available: https://www.piveau.io.

[12] SEMIC / EU Publications Office, "DCAT Application Profile for Data Portals in Europe (DCAT-AP)," version 3.0, [Online]. Available: https://semiceu.github.io/DCAT-AP/releases/3.0.0/

[13] International Data Spaces Association, "IDS Reference Architecture Model," version 4.2, [Online]. Available: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/