

DIGITAL-2023-CLOUD-AI-04-ICU-DATA

INDICATE

Deliverable **D3.1**

**ELSI-framework for data access, Data
Management Plan, and Data Protection
Impact Assessment:
ELSI-framework**

Cover Page

PROJECT INFORMATION	
Project number:	101167778
Project acronym:	INDICATE
Project name:	A federated INfrastructure for Data of Intensive CAre uniTs in Europe
Project starting date:	01-12-2024
Project duration:	42 months

DOCUMENT INFORMATION	
Deliverable:	D3.4 – ELSI-framework for data access, Data Management Plan, and Data Protection Impact Assessment
Work Package:	WP3 – Governance, Business Models and Legal Framework
Lead Contributor	Erasmus MC
Primary Author(s)	Jan van den Brand (Erasmus MC), Name (acronym beneficiary)
Dissemination Level ¹	PU
Deliverable Type ²	R
Contractual delivery date	31-5-2025
Actual delivery date	04-07-2025
Document status	Final

TEAM	
Authors:	Jan van den Brand (Erasmus MC)
Contributors:	Natalja Zilinski (UDUS) Celia Alvarez (SAS) Christine Riou (CHU Rennes) Viola Woeckel (Erasmus MC)
Acknowledgements:	Name (acronym beneficiary) Name (acronym beneficiary)
Reviewers:	Falk von Dincklage (UMG) Boris Delange (CHU Rennes)

HISTORY OF CHANGES			
Version	Date (DD-MM-YYYY)	Author/reviewer	Description
0.1	20-5-2025	Jan van den Brand	Initial version
0.2	17-6-2025	Jan van den Brand	Added Template Data Sharing Agreement
0.3	01-07-2025	Jan van den Brand	Adressed comments by Boris Delange (reviewer) Added: <ul style="list-style-type: none"> • Societal Implications section •

¹PU = Public; SEN = Sensitive, limited under the conditions of the Grant Agreement; CO = Confidential, only for members of the Consortium.

²R= Document/Report; DEC = Website; DEM = demonstrator; DATA = federated datasets

Table of Content

Cover Page	2
Introduction	5
Work accomplished	6
Legal Framework	10
Analysis of applicable regulations for INDICATE	10
Prioritized Regulatory Requirements	10
Key Regulatory Requirements for INDICATE	10
Contractual Framework for INDICATE	11
Constitutive Agreement	11
Accession Agreement	12
Data Sharing Agreement	12
Data Contract	12
Ethical Framework	13
Ethical Approvals for re-use of health data	13
Data Protection	13
Data Management	13
Discussion and conclusions	3
Appendices	15
Appendix 1: Regulatory Landscape	15
Appendix 2: Template Data Sharing Agreement	15
Appendix 3: Template Data Contract	15
Appendix 4: Data Protection Impact Assessment	16
Appendix 5: Data management plan	16

Introduction

INDICATE (INfrastructure for DIstributed and seCure Access To intEnsive care data) aims to establish a pan-European federated data infrastructure connecting Intensive Care Units (ICUs) across Europe. The key feature is that it enables secure data sharing and analytics for data users while ensuring that personal data remains under control of the data providers, addressing both innovation needs and regulatory compliance.

Federated data within INDICATE refers to an architectural pattern where personal data remains under the control of the Data Provider, while Data Users can request data processing in order to execute analytical queries, train machine learning models, or run machine learning models for inference. Personal data is defined under GDPR and for ease of interpretation considered synonymous with personally identifiable information as defined under HIPAA. However, INDICATE must first and foremost comply with GDPR. Personal data remains at the Data Provider either because it stays located physically at the data provider (on-premises or in private cloud) or under strict access control in the virtual private cloud environment of the Data Provider. Only aggregated results, model gradients, or model weights may be transferred to centralized systems or Data Users.

The implementation follows a carefully planned, phased approach spanning three years. The foundation phase begins in 2025 with the onboarding of three flagship academic medical centers, establishing core infrastructure and processes. The expansion phase in 2026 scales the network to nine institutions across Europe, validating the infrastructure's scalability and refining operational procedures. By the maturity phase in 2027, the network will encompass sixteen connected institutions, demonstrating full operational capability. Post-project growth continues with the addition of six new institutions annually, ensuring sustainable expansion of the network.

INDICATE's security framework addresses critical concerns through privacy-preserving computation that keeps sensitive data local. The zero-trust architecture, combined with end-to-end encryption and comprehensive audit trails, ensures robust security. The platform's design inherently supports GDPR compliance and aligns with European Health Data Space requirements while meeting NIS2 Directive specifications. This security-first approach maintains local data sovereignty while enabling collaborative analysis. The security design and architecture goes hand in hand with the Ethical and Legal framework and iterative improvement of both security and ELSI frameworks is envisioned through the development and operations of INDICATE.

The immediate path forward involves systematic stakeholder engagement and requirements gathering, followed by technical assessment and gap analysis. A pilot implementation with key departments precedes the phased rollout, which aligns with organizational priorities. Success metrics include quantitative measures such as the number of connected institutions and research collaborations, as well as qualitative assessments of clinical decision support adoption and operational efficiency improvements.

Work accomplished

The INDICATE project (A federated Infrastructure for Data of Intensive Care units in Europe) has made significant progress in establishing a comprehensive Ethical, Legal, and Social Implications (ELSI) framework for data access. This framework serves as the foundation for secure, compliant, and ethical sharing of ICU data across European healthcare institutions.

Legal Framework Development

Erasmus MC (Lead)

A thorough analysis of the European regulatory landscape for health data sharing has been conducted, resulting in a prioritized overview of regulatory requirements. This analysis identified key regulations affecting the INDICATE infrastructure, including GDPR, the Data Governance Act, Data Act, European Health Data Space Regulation, Digital Services Act, eIDAS Regulation, AI Act, and Medical Device Regulation.

The team has established a prioritization methodology using the MoSCoW approach (Must, Should, Could, Won't) to categorize requirements based on their criticality for compliance. This provides clear guidance for implementation teams on which aspects must be addressed immediately versus those that can be addressed in later development phases.

Key regulatory requirements have been translated into specific architectural principles and technical specifications, ensuring that privacy-by-design, data sovereignty, and lawful processing are embedded in the INDICATE infrastructure from the ground up.

Contractual Framework Implementation

Erasmus MC (lead), UDUS (contributor)

The project has identified a structured contractual framework comprising four key components:

- **Constitutive Agreement:** Currently based on the Consortium Agreement, with plans to develop a standalone agreement when INDICATE transitions to a legal entity under European Law. This will follow the SITRA Template Constitutive Agreement.
- **Accession Agreement:** Identified as a requirement for the second development plateau to facilitate onboarding of new organizations to INDICATE.
- **Data Sharing Agreement:** A template has been created to standardize bilateral agreements between Data Providers and Data Users, covering legal and organizational aspects of data sharing.
- **Data Contract:** A technical specification format (based on YAML) has been adopted to define the structure, semantics, quality, and terms of data exchange. This machine-readable approach enables automated validation and enforcement of data sharing rules.

Ethical Framework Development

Data Protection

Erasmus MC (lead), CHU Rennes (Contributor)

A comprehensive Data Protection Impact Assessment (DPIA) has been conducted for INDICATE's central services (the Hub). This assessment:

- Identified and evaluated potential privacy risks
- Established risk mitigation measures

- Created a monitoring framework for ongoing risk management
- Designated responsibilities for implementation and review

The DPIA has been structured according to the UK Information Commissioner's Office template, providing a detailed analysis of data processing activities, necessity and proportionality assessments, and specific measures to reduce identified risks.

Ethical Approval Processes

Erasmus MC (Lead), KPMG (Contributor), SAS (Contributor)

The project has developed an ethical framework that addresses:

- Ethical approval processes for health data re-use
- Data protection principles, with emphasis on Architecture Principle 15: "Personal Data should remain at the Data Provider"
- Data management planning for sensitive ICU patient data

Documentation Development

Erasmus MC (Lead), UDUS (Contributor), CHU Rennes (Contributor), KPMG (Contributor), SAS (Contributor)

Several critical documents have been created to support the ELSI framework:

- Regulatory Landscape Analysis: A comprehensive examination of applicable regulations
- Template Data Sharing Agreement: Standardizing bilateral data sharing terms
- Template Data Contract: Following Data Contract Specification 1.1.0 format for the MIMIC-EU dataset
- Data Protection Impact Assessment: For INDICATE's central services
- Data Management Plan: Outlining the data lifecycle within the INDICATE infrastructure

Governance Structure Implementation

Erasmus MC (Lead), KPMG (Contributor),

The first iteration on the framework establishes governance mechanisms for data access and sharing, including:

- Roles and responsibilities for Data Providers, Data Users, and the INDICATE governance authority
- Decision-making processes for data access requests.
- Compliance monitoring procedures.
- Dispute resolution mechanisms.

This comprehensive approach ensures that the INDICATE infrastructure will operate in full compliance with European regulations while facilitating valuable research access to ICU data through privacy-preserving federation mechanisms.

Societal Implications

To strengthen the social implications component of the ELSI framework, INDICATE will establish robust mechanisms for transparency and public engagement that go beyond the current stakeholder consultation processes. This enhancement includes the development of a public-

facing transparency communication materials that serve as a centralized view for all activities conducted within the INDICATE infrastructure. This would provide citizens with accessible information about, for example, active studies utilizing ICU data, including study purposes, participating institutions, privacy safeguards employed, and anticipated societal benefits. The communication will implement a tiered information architecture with summary views for general public consumption and detailed technical specifications for interested stakeholders. Additionally, INDICATE will establish structured public engagement mechanisms including annual public consultations, patient advisory panels with rotating membership, and accessible feedback channels that allow citizens to voice concerns or suggestions about data usage policies. These mechanisms will be complemented by proactive communication strategies that translate complex research outcomes into understandable public health insights, demonstrating the tangible benefits of ICU data sharing for European healthcare improvement.

Discussion and conclusions

The development of the ELSI framework for INDICATE marks a significant milestone in establishing governance for a pan-European federated ICU data infrastructure. This framework embodies privacy-by-design principles, ensuring personal health data remains under the control of data providers while only sharing aggregated results, addressing legal and ethical concerns around patient privacy. This approach supports the European Health Data Space objectives and respects member state sovereignty over health data governance. The regulatory landscape analysis highlights the complexity of European regulations affecting health data sharing, with INDICATE's prioritization methodology providing a practical framework for implementation teams to navigate this complexity and ensure compliance. The contractual framework, built on established data space principles, demonstrates INDICATE's commitment to interoperability with broader European data space initiatives, positioning it as a potential reference implementation for health domain-specific applications. The ELSI framework's comprehensive approach is expected to positively impact the project's results by providing clear guidance for technical implementation teams, reducing compliance gaps, and accelerating the onboarding of new data providers through standardized mechanisms. The emphasis on data sovereignty and patient privacy aligns with healthcare institutions' concerns, reducing resistance to participation. The framework demonstrates that cross-border health data sharing can respect both scientific needs and ethical imperatives. The adoption of machine-readable data contracts alongside traditional legal agreements bridges the gap between legal and technical domains, improving scalability and reliability of compliance.

Recommendations for future work include further standardization of the Data Contract framework, systematic monitoring and updating of the ELSI framework, guidance for Data Providers on implementing compliant local components, expansion of the ethical framework to address AI in healthcare, and alignment with emerging European enforcement frameworks for data spaces. These recommendations highlight the need for INDICATE to maintain an adaptive approach as the regulatory landscape and technical capabilities evolve.

Supplements: Legal Framework

Analysis of applicable regulations for INDICATE

The INDICATE project (A federated INfrastructure for Data of Intensive CAre uniTs in Europe) aims to establish a pan-European federated data infrastructure for secure cross-border access to Intensive Care Unit (ICU) datasets. This infrastructure must navigate a complex regulatory landscape encompassing data protection, privacy, security, and healthcare-specific regulations.

An extensive analysis of the regulatory landscape for sharing health data in EU was performed to inform the design of INDICATE. This resulted in a prioritized overview of regulatory requirements. The overview is listed in Appendix 1 and will inform the continued development of INDICATE.

Prioritized Regulatory Requirements

The document prioritizes regulatory requirements using the MoSCoW method (Must, Should, Could, Will not) and ranks the applicable regulations in descending order of priority:

- **General Data Protection Regulation (GDPR)** - The foundation of all data protection efforts, governing personal data processing, data subject rights, and security measures.
- **Data Governance Act (DGA)** - Establishes frameworks for data sharing and governance, ensuring secure and efficient data access across sectors and borders.
- **Data Act** - Promotes data availability for business and research use, reducing barriers to data access and enhancing interoperability.
- **European Health Data Space Regulation (EHDS)** - Specifically focused on health data, ensuring individual control and enabling research and policy-making.
- **Digital Services Act (DSA)** - Sets standards for transparency, accountability, and user protection for online platforms and digital services.
- **eIDAS Regulation** - Provides a framework for secure electronic identification and trust services.
- **AI Act** - Establishes requirements for safe and ethical AI development and use, including risk management and data governance.
- **Medical Device Regulation (MDR)** - Sets safety and performance requirements for medical device software.

Key Regulatory Requirements for INDICATE

The infrastructure must implement measures to address:

- **Data Minimization and Sovereignty** - Only essential data should be collected and processed, with personal data remaining under the control of the data provider. In addition, data and services will be hosted only in EU Member States and aligned with EU data protections and cybersecurity standards. Finally, to ensure freedom of jurisdiction (e.g. from the US Cloud Act) fall back scenarios have been created to ensure that key data will be stored in a different location governed by an EU entity.
- **Lawful Processing Basis** - Clear legal grounds for processing both user data and patient data must be established.

- **Data Subject Rights** - Mechanisms for individuals to exercise their rights to access, rectify, erase, and restrict processing must be provided.
- **Security by Design** - Robust security measures including encryption, access controls, and auditing capabilities must be implemented. Furthermore, standard operating procedures are to be developed in a follow-up task that detail the audit procedures (T3.2 Standard Operating Procedures).
- **Incident Reporting** - Procedures for reporting data breaches and cybersecurity incidents within regulatory timeframes (GDPR: 72 hours; NIS2: 24 hours initial notification).
- **Cross-Border Data Sharing** - Standardized data transfer agreements and appropriate safeguards for international data transfers must be in place.
- **Data Quality and Validation** - Mechanisms to ensure data quality, accuracy, and reliability must be implemented.
- **Governance and Accountability** - Comprehensive documentation, impact assessments, and appointment of responsible roles (such as a Data Protection Officer) are required.

These regulatory requirements shape INDICATE's architecture, governance model, and operational procedures, ensuring that the federated data infrastructure can deliver value while maintaining compliance with the European legal framework for data sharing and protection.

The requirements have been further specified in the INDICATE and listed in the INDICATE requirements repository. They have been incorporated into the Architecture of INDICATE under the responsibility of the Technical Lead. A review process for the implementation of the design has been established with a formal Architecture Review Board (part of the External Expert Advisory Committee).

Contractual Framework for INDICATE

The Contractual Framework of INDICATE encompasses all agreements that govern the operation and activities of INDICATE and its participants. It establishes relationships, contractual rights and duties, implements the data space governance framework, and makes part of its rules legally enforceable among data space participants. Beyond data space governance, the Contractual Framework also enables data space participants, particularly data providers and, indirectly, data rights holders, to exercise their data sovereignty and add conditions to the baseline governance provided by the data space governing authority.

The Contractual Framework of INDICATE is based upon the SITRA Rulebook for a Fair Data Economy¹ and the Data Space Support Center's guidance on Contractual Framework for data spaces.²

Constitutive Agreement

The Constitutive Agreement underlies the creation of INDICATE as a data space open to participation and provides a minimum mandatory governance framework at the data space level applicable to all participants. It sets out rules that regulate the relationship between all data space participants, binding them to the INDICATE governance framework.

¹ <https://www.sitra.fi/en/publications/rulebook-for-a-fair-data-economy/#2-3-introduction-to-rulebook-contractual-framework>

² <https://dssc.eu/space/BVE2/1071254557/Contractual+Framework>

At the start of the grant period, the Consortium Agreement between the INDICATE participants functions as the Constitutive Agreement. The aim of INDICATE is to lay the groundwork for a legal entity under European Law. Upon the creation of this entity a renewed Constitutive Agreement will be signed. This Constitutive Agreement will follow the Template Constitutive Agreement created by SITRA.³ It will describe:

- General Terms and Conditions
- Governance Model
- Description of the ICU data space
- Accession Agreement

As recommended by SITRA, the general terms and conditions document will not be amended. Instead, any relevant amendments will be added as derogations from the constitutive agreement. This will allow members to easily identify what amendments have been made without comparing to the original terms and conditions.

Accession Agreement

INDICATE will facilitate organizations, including Data Providers, Data Users and Service Providers to join the data space. To that end, an Accession Agreement will be required to lay out rights and responsibilities for all participants in INDICATE. The Accession Agreement will be created during the second development plateau of INDICATE and is currently out of scope.

Data Sharing Agreement

The Data Sharing Agreement is a bilateral agreement between the Data Provider and Data User that describes the Terms and Conditions for data access, including the processing purpose, limitation of processing and appropriate legal base for processing of personal data. The Data Sharing Agreement covers the legal and organizational aspects of Data Sharing in INDICATE and is supported by the Data Contract.

INDICATE has created a Template Data Sharing Agreement that can be used by Data Providers to create a Data Sharing Agreement. The template is listed in Appendix 2.

Data Contract

Bilateral agreement that defines the structure, format, semantics, quality, and terms of use for exchanging data between a data provider and their consumers. Think of an Application Programming Interface, but for data. A data contract is implemented by a data product offered by a Data Provider. Data contracts can also be used for the input port to specify the expectations of data dependencies and verify given guarantees.⁴ In other words, it is a technical and both human and machine-readable specification to support secure data access to high quality data.

The data contract specification defines a YAML format to describe attributes of provided data sets. It is data platform neutral and can be used with any data platform, such as AWS S3, Google BigQuery, Azure, Databricks, and Snowflake. The data contract specification is an open initiative to define a common data contract format. It follows OpenAPI and AsyncAPI conventions. There are several tools available, both commercial and non-commercial open source to assist in the specification and review of data contracts.

³ <https://www.sitra.fi/wp/wp-content/uploads/2025/03/rulebook-model-for-a-fair-data-economy-part-2-v3d.pdf>

⁴ <https://datacontract.com/>

Data contracts come into play when data is shared between Data Provider and Data User. First, and foremost, data contracts are a communication tool to express a common understanding of how data should be structured and interpreted. They make semantic and quality expectations explicit. They are often created collaboratively in workshops together with data providers and data consumers. Later in development and production, they also serve as the basis for code generation, testing, schema validations, quality checks, monitoring, access control, and computational governance policies.

The creation of a data contract is supported by the Free Open Source Software available from <https://datacontract.com/>. A template Data Contract for MIMIC-EU (WP6, D6.1) is include in Appendix 3.

Ethical Framework

Ethical Approvals for re-use of health data

It is the responsibility of the Data User to obtain relevant ethical approvals or waivers for the re-use of health data, as a one size fits all approach is not feasible nor desirable for INDICATE. After all, it is the Data User who is best suited to describe the purpose and scope of processing required to meet their goals. Next, it is up to the Data Provider to ensure that the approval or waiver meets the Terms and Conditions for data sharing. For example, Terms and Condition include the legal base for data processing and if processing for commercial purposes is allowed.

Data Protection

Data sovereignty and trust is one of the core pillars for INDICATE. The key guiding principle for INDICATE is Architecture Principle 15: Personal Data should remain at the Data Provider to ensure privacy-by-design-and-default, and data sovereignty. This principle guides import design decisions. For example, as a consequence INDICATE follows a Hub-Spoke Architecture where a limited set of centralized services enables secure connectivity and transfer of aggregated data between the data providers and data users. The central hub will be hosted on Microsoft Azure in a EU Member State (either West Europe [The Netherlands] or North Europe[Ireland])

A Data Protection Impact Assessment (DPIA) for INDICATE's central services (the Hub) has been performed. The spokes have not been included in the DPIA, since INDICATE has no control over the data that resides at the Data Provider and it is the responsibility of the individual Data Provider to perform a DPIA the spoke services that they deploy in support of connectivity to INDICATE.

The DPIA is included in Appendix 4.

Data Management

The INDICATE Data Management Plan outlines a comprehensive framework for handling sensitive ICU patient data across European healthcare institutions through a federated infrastructure approach. The plan emphasizes that patient data will never leave hospital control, with analysis tools being brought to the data rather than extracting data for external analysis. This privacy-by-design approach ensures compliance with GDPR and other relevant regulations while still enabling cross-border research collaboration. The DMP details how data will be standardized using the OMOP Common Data Model and HL7 FHIR standards, with rich metadata following HealthDCAT-AP specifications to enhance

discoverability. It addresses the FAIR principles, establishing protocols for making data findable and interoperable through standardized vocabularies like SNOMED CT and LOINC. The plan also covers data security measures, ethical considerations, and resource allocation, while emphasizing that the federated architecture eliminates the need for centralized data repositories and instead relies on secure, controlled access mechanisms where only aggregated, non-identifiable results are shared between institutions.

The Data Management Plan is included in Appendix 5.

Appendices

It is possible to add supplemental information to this deliverable, but please note these will be merged into one file and shared with the reviewers, EMT, and the EC. Deliverables will be accessible by consortium members, but not to the public (unless indicated as such in the dissemination level). If you wish to refer to other relevant information sources, such as references or other projects/websites, please list them here.

Appendix 1: Regulatory Landscape

[INDICATE RegulatoryLandscape_ErasmusMC_v1.0_20250403.docx](#)

Appendix 2: Template Data Sharing Agreement

[INDICATE TemplateDataSharingAgreement_20250613_v0.2.docx](#)

Appendix 3: Template Data Contract

This template data contract for the MIMIC-EU dataset follows the Data Contract Specification 1.1.0 format and includes:

1. Basic Information

- A unique identifier (urn:datacontract:indicate:mimic-eu)
- Metadata including title, version, description, ownership, and contact details
- Status set as "in development"

2. Server Information

- Production and staging environments using S3 storage
- Role-based access control for different research purposes
- File format specifications (Parquet)

3. Terms and Conditions

- Clear usage guidelines and limitations for research purposes
- Privacy and ethics policies
- Billing and notice period information

4. Data Models

- OMOP CDM tables
 - person
 - observation_period
 - death
- visit_occurrence
- condition_occurrence
- drug_exposure
- procedure_occurrence
- measurement
- observation

5. Data Quality Rules

Each model includes SQL-based quality rules to ensure:

- Proper record counts

- Valid date ranges
- Logical consistency in sequences (e.g., discharge after admission)
- Completeness of required fields

6. Service Level Agreements

- Availability of the data: 99.5% of the time
- Data retention: 5 years
- Latency: 72 hours maximum
- Data update frequency: Monthly
- Support hours and response times
- Backup procedures

The YAML file includes the full example specification of the Data Contract for MIMIC-EU demonstrator

[data-contract-mimic-eu.yaml](#)

Appendix 4: Data Protection Impact Assessment

[INDICATE DpiaIndicateCentralServices_ErasmusMC_v0.2_20250602.docx](#)

Appendix 5: Data management plan

[INDICATE Deliverable D2.1 DataManagementPlan_SAS_v0.2_20250429.docx](#)

