DARE

DIGITAL LIFELONG PREVENTION

CODE NO. PNC0000002

Spoke 1 Deliverable

S1.D3.2 Multidimensional interoperability strategies and solutions deployed

This research is co-funded by the Ministry of University and Research within the Complementary National Plan PNC-I.1 "Research initiatives for innovative technologies and pathways in the health and welfare sector"
D.D. 931 of 06/06/2022, PNC0000002 DARE - Digital Lifelong Prevention

Ministero dell'Università e della Ricerca | PNC | DARE

S1.D3.2 Multidimensional interoperability strategies and solutions deployed

Deliverable information	
Spoke number and title	Spoke 1 - ENABLING FACTORS AND TECHNOLOGIES FOR DIGITAL PREVENTION
WP number and title	WP 3 - Interoperability Governance
Related task(s)	Task 3.2 - Building Trust
Lead beneficiary	INFN
Dissemination level	Public
Due date	14/12/2025
Actual date of delivery	14/12/2025
Author(s)	Barbara Martelli (INFN)
Contributors	Francesco Sinisi, Giusy Sergi, Stefano Zotti, Alessandro Costantini, Andrea Chierici, Mauro Patano, G. Donvito, C. Vistoli (INfN)
Quality Assurance	Marco Viceconti, Sabato Mellone (UNIBO)

Document history

Version	Date	Author(s) /Reviewer(s) (Beneficiary)	Description
0.1	02/10/2025	Barbara Martelli (INFN)	First draft
0.2	17/10/2025	Barbara Martelli	Additions to Sections 2 and 4
0.3	08/11/2025	Barbara Martelli	Section 3
0.4	10/11/2025	F. Sinisi, G. Sergi, S. Zotti, A. Costantini, A. Chierici, M. Patano, G. Donvito, M.C. Vistoli (INFN)	Integrations
0.5	14/12/2025	M. Viceconti, S. Mellone	Revision
1.0	14/12/2025	Barbara Martelli	Final document

Disclaimer

Table of contents

Publishable summary

This deliverable presents the strategies and solutions for multi-dimensional interoperability deployed within Spoke 1 of the DARE project, focusing on the activities of WP3. The document details the governance, technical, semantic, organizational, and legal interoperability frameworks established to enable seamless data exchange and integration across the DARE ecosystem, supporting pilots in Spoke 2 and Spoke 3. The deliverable builds upon the requirements and architecture defined in D3.1 and incorporates lessons learned from the deployment and early operation of interoperability solutions in real-world settings.

1. Introduction

1.1. Purpose and Scope

The purpose of D3.2 is to document the strategies, methodologies, and concrete solutions adopted and deployed to achieve multi-dimensional interoperability within the DARE project. The focus is on Spoke 1's role as a solution provider, supporting pilots in Spoke 2 (community-based digital primary prevention) and Spoke 3 (digitally-enabled secondary and tertiary prevention). This deliverable addresses technical, semantic, organizational, and legal interoperability, in line with national and European standards and the FAIR principles.

1.2. Relation to Other Deliverables

D3.2 builds directly on D3.1 ("Analysis of the technological framework and interoperability requirements"), which defined the initial architecture and requirements. It also aligns with deliverables from WP2 (legal and ethical framework), WP4 (technology and analytics), and WP5 (impact analysis and upscaling), ensuring a holistic approach to interoperability.

2. Background and Context

2.1. The DARE Project and Spoke 1

DARE (DigitAl lifelong pRevEntion) is a national initiative to transform healthcare in Italy through digital prevention, leveraging data-driven approaches and advanced digital technologies. Spoke 1 acts as the competence center for enabling factors and technologies, providing solutions and governance for interoperability, data management, and technology transfer.

2.2. Interoperability in DARE: Vision and Challenges

Interoperability is recognized as a critical enabler for the DARE ecosystem, addressing the fragmentation of health information systems and supporting the integration of heterogeneous data sources (clinical, environmental, behavioral, etc.). The main challenges include:

- Heterogeneity of data formats and standards

- Legacy systems and proprietary solutions
- Legal and ethical constraints (GDPR, data protection)
- Organizational silos and stakeholder alignment
- Need for scalability and sustainability

3. Multi-dimensional Interoperability Strategy

3.1. Governance and Coordination

The governance and coordination model adopted in DARE is multi-layered and participatory, designed to ensure effective alignment between enabling technologies (Spoke 1) and the real-world needs and constraints of pilot implementations (Spoke 2 and Spoke 3). This model is built on the principles of transparency, inclusiveness, and adaptability, and is structured to support both top-down strategic direction and bottom-up innovation from pilot projects and Solution Frameworks.

A multi-dimensional governance model has been established, involving:

- Direct links with pilot sites in Spoke 2 and Spoke 3. A key feature of the governance model is the establishment of direct, bidirectional links between Spoke 1 and the pilot sites in Spoke 2 and Spoke 3. Each pilot is assigned a dedicated liaison team responsible for:
    - Translating pilot-specific needs into technical and organizational requirements.
    - Facilitating the adoption and customization of interoperability solutions.
    - Ensuring feedback from pilot implementation is rapidly integrated into the evolution of the overall architecture.

- Spoke2 and Spoke3 pilots as Living Labs: Spoke2 and Spoke3 pilots serve as living laboratories for testing and refining interoperability strategies. They provide a diverse set of real-world scenarios—ranging from community-based digital prevention to integration of environmental, behavioral, and clinical data—allowing the governance model to be stress-tested and iteratively improved. Regular coordination meetings, joint workshops, and shared documentation platforms foster continuous dialogue and knowledge exchange between technology providers and pilot implementers.
- Coordination with external providers and legacy systems: the governance framework also includes mechanisms for coordinating with external service

Commentato [SM1]: Perché solo spoke 2? È una svista o è intenzionale?

Commentato [BM2R1]: Svista, corretto

providers and integrating legacy systems. This is achieved through formal agreements, shared technical guidelines, and the deployment of middleware and APIs that bridge heterogeneous infrastructures. Examples of this are integrations of Ansys products into INFN EPIC Cloud.

- Continuous alignement and adaptation with national and European interoperability frameworks such as EHDS, EOSC: the governance model is designed to be adaptive, with regular reviews and updates based on lessons learned from pilot deployments and changes in regulatory or technological landscapes. Cross-spoke working groups and data governance committees are empowered to propose modifications, ensuring that the model remains responsive and future-proof. This is integrated in EPIC Cloud processes.

Commentato [SM3]: Questi, e possibilmente altri casi, conviene descriverli come esempi applicativi nella sezione degli use case.

Commentato [BM4R3]: Per me e' ok, da capire pero' se c'e' tempo (e' necessario che arrivi un contributo dai responsabili degli use case che non sono di INFN), @sabato riesci a coinvolgerli tu?

Commentato [SM5R3]: Scrivo nelle sezioni successive un po' di informazioni raccolte sui piloti. Qui chiuderei cosi.

3.2. Technical Interoperability

Technical interoperability in DARE is achieved through a layered, standards-based approach that ensures seamless, secure, and scalable data exchange across a heterogeneous ecosystem of health, research, and public sector infrastructures. The strategy is designed to address the complexity of integrating diverse data sources, legacy systems, and emerging digital heal

Key actions and solutions include:

- **Adoption of open standards for data exchange**: DARE prioritizes the use of internationally recognized standards for health data exchange and representation. Key standards include:
  - **HL7 FHIR** for structured health data interoperability, enabling granular, RESTful access to clinical information.
  - **DICOM** for medical imaging, ensuring compatibility with radiology and imaging systems.
  - **OMOP CDM** for harmonizing observational health data and supporting secondary use in research and analytics.
  - **SNOMED CT, LOINC, ICD** for semantic consistency in clinical terminologies.
- **Deployment of middleware and APIs to bridge heterogeneous systems**: to bridge the gap between heterogeneous systems, DARE deploys middleware solutions and standardized APIs. These components enable:
  - Real-time and batch data exchange between clinical, environmental, behavioral, and research data sources.
  - Secure, federated access to distributed datasets, supporting both centralized and edge-computing scenarios.

  - Integration with legacy hospital information systems and external data providers through adapters and connectors.
- Use of Salus Ratio (Solution Framework #2), develoed using AlmaHealthDB as a reference architecture, for secure, scalable data integration. The technical architecture is modular and service-oriented. Key features include:
  - **Cloud-native applications** deployed on INFN Cloud, CINECA HPC, ReCaS DataCenter, EOSC, and IRCCS datacenters, ensuring portability and elasticity.
  - **Containerization and orchestration** (e.g., Kubernetes, RKE2) for reproducible deployment and management of services across multiple infrastructures.
  - **Federated identity and access management** using solutions like Keycloak and FreeIPA, supporting single sign-on, multi-factor authentication, and fine-grained authorization.
- **Data Management and Security:** robust data management practices are enforced to guarantee data integrity, confidentiality, and traceability:
  - **RUCIO/FTS** is used for secure, policy-driven data movement and replication, supporting large-scale, distributed data workflows.
  - **Encryption at rest and in transit**, network segmentation, and audit logging are implemented to meet GDPR and national data protection requirements.
  - **Consent management frameworks** and data access policies are integrated to ensure compliance with ethical and legal standards.
- **Interoperability Testing and Continuous Integration: t**echnical interoperability is validated through continuous integration pipelines, automated testing, and pilot deployments. This includes: end-to-end testing of data flows between pilot sites and core infrastructure; monitoring and alerting for data quality, system performance, and security incidents; iterative refinement based on feedback from Spoke 2 and Spoke 3 pilots, ensuring that solutions remain robust and responsive to evolving needs.
- **Scalability and Future-Proofing:** the architecture is designed to be scalable and adaptable, supporting the onboarding of new data sources, partners, and use cases. Continuous alignment with evolving standards (e.g., updates to HL7 FHIR, OMOP CDM) and emerging technologies (e.g., AI/ML integration, edge computing) ensures that DARE remains at the forefront of technical interoperability in digital health.
- **Implementation of cloud-native applications and containerization for portability across infrastructures** INFN Cloud, CINECA HPC, ReCaS DataCenter, EOSC, Research Hospitals (IRCCS) datacenters).

3.3. Semantic Interoperability

- Mapping and harmonization of terminologies (ICD, SNOMED, LOINC)
- Use of semantic converters (e.g., FHIR-to-OMOP CDM) to enable secondary use of health data
- Ontology-driven data integration for research and analytics
- FAIR-by-design approach to ensure data findability, accessibility, interoperability, and reusability

Semantic interoperability is a foundational pillar in DARE, ensuring that data exchanged across diverse systems and organizations is not only syntactically compatible but also carries consistent meaning. This capability is critical for enabling integrated analytics, secondary use of health data, and collaborative research across the DARE ecosystem, especially in the context of Spoke 2 and Spoke 3 pilots.

Standardized Terminologies and Ontologies: DARE adopts internationally recognized clinical terminologies and coding systems to harmonize the representation of health concepts:

- **SNOMED CT** for comprehensive clinical terminology, supporting detailed and unambiguous encoding of diagnoses, procedures, and findings.
- **LOINC** for laboratory and clinical observations, enabling standardized reporting of test results and measurements.
- **ICD (International Classification of Diseases)** for disease classification and epidemiological studies.
- **ATC (Anatomical Therapeutic Chemical Classification)** for medications and treatments.

Semantic Mapping and Harmonization: to bridge the gap between heterogeneous data sources, DARE implements semantic mapping strategies:

- **Terminology mapping tools** are used in Salus Ratio to translate local codes and legacy vocabularies into standard terminologies, ensuring data consistency across sites.
- **Ontology-driven data integration** leverages domain ontologies to align disparate data models, supporting advanced queries and cross-domain analytics.
- **FHIR-to-OMOP CDM converters** enable the transformation of clinical data from HL7 FHIR resources into the OMOP Common Data Model, facilitating secondary use for research and population health studies.

Commentato [BM6]: @Sabato controlla per favore questa parte

Commentato [SM7R6]: Tutti gli standard citati sono stati effettivamente impiegati almeno 1 volta.

Metadata and FAIR Principles: semantic interoperability is reinforced by the adoption of FAIR (Findable, Accessible, Interoperable, Reusable) data principles:

- **Rich metadata schemas** are applied to all datasets, capturing provenance, context, and usage rights.
- **Persistent identifiers** (e.g., DOIs, UUIDs) are assigned to curated datasets and data elements, supporting traceability and reproducibility.
- **Data catalogs** are maintained in Zenodo **Data ingestion APIs:** DARE deploys services and APIs that provide **semantic annotation tools** to enrich unstructured data (e.g., clinical notes, imaging reports) with standardized concepts.

Continuous Alignment and Community Engagement: semantic interoperability is maintained through:

- **Regular updates** to terminologies and mappings, in line with evolving standards and clinical practice.
- **Collaboration with national and European initiatives** (e.g., EOSC, Health Big Data, ICSC, ELIXIR, BBMRI-ERIC, EOSC-Life, TEHDAS2, EUSAiR) to ensure alignment and reuse of semantic assets.
- **Training and capacity-building** for data stewards, clinicians, and IT staff to promote best practices in semantic data management.

Impact on DARE Pilots: these strategies enable DARE pilots to integrate and analyze data from multiple sources with semantic consistency; support advanced use cases such as digital twins, risk stratification, and AI-driven decision support; facilitate cross-site research and federated analytics, accelerating innovation in digital prevention and personalized healthcare.

Commentato [BM8]: @Sabato anche qui verifica per favore che sia davvero coerente con quanto offerto da AlmaHealthDB

Commentato [SM9R8]: Modificato

Commentato [BM10]: @Sabato aggiungi altre di vostra visibilita'

Commentato [SM11R10]: Fatto

3.4. Organizational Interoperability

Organizational interoperability in DARE refers to the alignment of processes, roles, and collaborative practices across all participating pilots, partners, and stakeholders. It is the foundation that enables diverse organizations—each with their own structures, cultures, and operational models—to work together effectively toward common goals in digital health innovation.

This dimension of interoperability is realized through several coordinated actions:

- **Definition of Shared Workflows and Protocols:** common workflows and standardized protocols are co-designed and adopted across pilots and partners. This ensures that data collection, processing, sharing, and usage follow harmonized procedures, reducing ambiguity and fostering trust among all actors.

- **Training and Capacity-Building for End-Users and IT Staff:** continuous education and skill development initiatives are implemented to empower all participants—clinicians, researchers, IT professionals, and data managers—to effectively use and contribute to interoperable systems. This includes targeted training sessions, workshops, and the development of comprehensive documentation.

By integrating these elements, DARE creates an organizational environment where interoperability is not just a technical achievement, but a shared commitment embedded in the daily operations and strategic vision of all partners. This approach is essential for scaling digital health solutions, sustaining innovation, and maximizing the impact of the DARE project across the national healthcare landscape.

Organizational interoperability is pursued also through the link that INFN provides to other communities like BBMRI-ERIC, EOSC Italian node, EGI.

Commentato [SM12]: A meno che non l'abbiate implementato voi in INFN per qualche pilota non credo vada riportato. Anche a livello di Fondazione ci sono state a limite azioni di comunicazione che hanno toccato l'interoperabilità ma nulla di simile a quello che veniva descritto qui.

Commentato [BM13R12]: In INFN facciamo queste cose nell'ambito del nostro User Forum, ma in effetti gli utenti DARE non sono coinvolti attualmente in questo forum (se si vorra' fare in futuro, si potranno coinvolgere). Vedi tu, se allo stato attuale preferisci toglierlo, per me e' ok

3.5. Legal and Ethical Interoperability

Legal and ethical interoperability is a critical dimension in the DARE project, ensuring that data sharing, integration, and processing across the ecosystem are not only technically feasible but also fully compliant with applicable laws, regulations, and ethical standards. In the context of digital health and biomedical research, this means that all activities must respect the rights and expectations of individuals, safeguard sensitive information, and adhere to national and European frameworks such as the GDPR, EHDS, and sector-specific codes of conduct.

Achieving legal and ethical interoperability requires a coordinated approach that goes beyond technical safeguards. It involves the alignment of policies, procedures, and governance structures across all partners and pilots, as well as the implementation of mechanisms for transparency, accountability, and continuous risk assessment. This includes:

- The adoption of shared legal frameworks and data protection agreements that define roles, responsibilities, and data usage conditions for all stakeholders.
- The establishment of ethical oversight committees and regulatory sandboxes, enabling the controlled experimentation and validation of innovative solutions—such as AI applied to medical data—within a compliant and ethically sound environment.
- The integration of consent management, data minimization, anonymization, and auditability into all data workflows, ensuring that individuals' rights are protected at every stage.
- Ongoing training and awareness-raising activities for researchers, clinicians, and technical staff to foster a culture of compliance and ethical responsibility.

By embedding legal and ethical considerations into the core of its interoperability strategy, DARE not only meets regulatory requirements but also builds trust among participants and the wider public, paving the way for responsible innovation and sustainable impact in digital health.

3.5.1. INFN EPIC Cloud as a concrete means to guarantee compliance

EPIC Cloud is a foundational asset for DARE to ensure compliance with GDPR and national data protection laws and to address ethical, privacy, and civil liability issues in collaboration with WP2. It's technological architecture and robust governance framework further enhance reliability, transparency, and trust for all stakeholders.

INFN EPIC Cloud is a dedicated region of the INFN DataCloud infrastructure, specifically designed to support the secure and compliant processing of sensitive biomedical and genomic data. It is certified under ISO/IEC 27001, 27017, and 27018, which are internationally recognized standards for information security and cloud privacy. This certification provides assurance that the platform's technical and organizational measures align with GDPR requirements and other relevant legal frameworks for health research.

Key compliance features include:

- **Strict Data Governance:** EPIC Cloud enforces policy-aware orchestration, ensuring that resource allocation and data processing respect legal and ethical frameworks such as GDPR, EHDS, and the Italian Data Protection Code. All processing activities are registered, risk-assessed, and lawfully conducted, with clear roles and responsibilities for Data Protection Officers and Security Groups.
- **Technical Safeguards:** The platform implements multifactor authentication, fine-grained audit logs, segregation of duties, IP whitelisting, encryption at rest and in transit, and network segregation between tenants. These controls are designed to protect personal and sensitive data, including genomic information that cannot be fully anonymized.
- **Organizational Controls:** EPIC Cloud maintains documented and tested disaster recovery and business continuity plans, redundant processing facilities for high availability, and regular reviews of compliance policies. Violations trigger disciplinary, contractual, or legal consequences, ensuring accountability.
- **Cloud Usage Policy:** Only authorized cloud services are permitted, with all services registered in the EPIC asset inventory. Sensitive workloads must apply encryption, segmentation, and verified identity controls. Provider compliance is monitored through periodic audits and certification reviews.
- **Support for Regulatory Sandboxes:** EPIC Cloud's architecture enables the creation of regulatory sandboxes for AI applied to medical data, allowing researchers

to experiment with new solutions in a controlled, compliant environment. This supports ethical innovation while ensuring legal requirements are met. Through ICSC, EPIC Cloud follows the the EUSAIR project (https://eusair-project.eu/) which offers a valuable opportunity to advance legal and ethical interoperability through the creation and operation of regulatory sandboxes specifically designed for AI applications in medical data. These sandboxes provide controlled environments where new AI-driven solutions can be developed, tested, and evaluated in compliance with relevant legal and ethical requirements.

- By leveraging EUSAiR's outputs in multi-stakeholder collaboration and innovation management, DARE can establish regulatory sandboxes that allow researchers, clinicians, and technology providers to experiment with AI models on sensitive medical datasets. These sandboxes facilitate the assessment of compliance with GDPR, EHDS, and other applicable regulations, while also enabling the evaluation of ethical implications such as patient consent, data anonymization, and algorithmic transparency. The sandbox approach enables rapid prototyping and validation of AI solutions, reducing the time and risk associated with bringing new technologies into clinical practice. It also fosters dialogue between regulators, researchers, and end-users, ensuring that legal and ethical considerations are addressed from the outset.

Through these synergies, INFN supports DARE objectives allowing the project to more effectively achieve its goals of federated, FAIR-compliant, and ethically sound data management for life sciences. This collaboration ensures that innovative AI applications are not only technically robust but also legally and ethically trustworthy, paving the way for their adoption in real-world healthcare settings.

3.5.2. Dedicated services implmeneted within Salus Ratio

Salus Ratio, in addition to the Governance and Technical Safeguards described for the EPIC Cloud, integrates two services: a synthetic data generation service—currently limited to structured data—and a service that instantiates sandboxes for the development and testing of algorithms. The latter service makes it possible to provide isolated environments tailored to developers' needs, where interoperability testing, adaptation of workflows to the data format managed within the infrastructure, and workflow development and deployment can be performed. Workflows developed and validated within the sandboxes can be imported by Salus Ratio's system administrators into the infrastructure for the processing of sensitive data and/or for the validation of the workflow/computational model using real data.

These two services play a crucial role in ensuring compliance with legal requirements for handling sensitive personal data. The synthetic data generation service enables early-stage development and testing without exposing real data, thereby reducing privacy risks.

Meanwhile, the sandbox environments provide controlled, isolated spaces where workflows and algorithms can be validated safely before being deployed on sensitive datasets. Together, they help maintain high standards of data protection and support lawful, secure processing within the infrastructure.

4.Solutions Deployed

4.1. Interoperability Architecture

The interoperability architecture deployed in DARE is designed to provide a secure, scalable, and standards-based foundation for multidimensional data integration and sharing across the national digital health ecosystem. At its core is the EPIC Cloud multisite architecture [https://doi.org/10.1051/epjconf/202533701113], which enables high availability, disaster recovery, and geographic distribution of sensitive biomedical and clinical data. This infrastructure supports the deployment of advanced data access and consent management frameworks, as documented in recent literature (see [DOI:10.22323/1.434.0028]), ensuring that all data processing activities are both legally compliant and ethically sound.

To further strengthen trust and compliance, DARE has developed comprehensive guidelines for secure data sharing and the secondary use of health data. These guidelines are operationalized through a modular, service-oriented architecture based on open standards, which facilitates the integration of diverse pilot sites via secure APIs and federated data access mechanisms.

The architecture's flexibility is demonstrated through a range of use cases—including multi-centric AI validation, federated data lakes, and medical imaging platforms—that showcase its ability to support ethical and legal compliance at scale. By leveraging both cloud and edge computing resources, the DARE interoperability architecture ensures that data processing is not only robust and efficient, but also adaptable to the evolving needs of research and clinical practice.

4.2. Data Flows and Integration

Some key achievements of WP3 have been reached by INFN through the evolution of the RUCIO/FTS data management system in sinergy with ICSC Spoke8. A PoC of a hardened RUCIO/FTS system has been deployed in the context of ICSC and is now ready to be deployed in a production environment on INFN EPIC Cloud. The RUCIO/FTS data management system is a cornerstone for secure, scalable, and compliant data management in biomedical research, in particular it is an enabler for:

- Implementation of data pipelines for structured and unstructured data

- Real-time and batch data exchange between clinical, environmental, and research systems
- Data curation and quality assurance processes in line with FAIR principles

The RUCIO PoC demonstrates the functionality of RUCIO as an open-source, secure Data Management for Life Sciences sensitive data management, anonymized data exchange, and continuous security monitoring, supporting FAIR and Open Science principles. It is worth mentioning the following aspects:

- Integration with Keycloak (authentication/authorization) and FTS (file transfer service) creates a robust ecosystem for managing sensitive health data.
- GDPR Compliance: The architecture is designed to ensure confidentiality, integrity, and availability of medical and omics data, with a strong focus on security hardening and continuous vulnerability assessment.
- Distributed Data Management: RUCIO coordinates data replication and placement, FTS handles large-scale transfers, and Keycloak manages access control, enabling secure handling of petabytes to exabytes of data.

4.3. Pilot Use Cases

The DARE project's approach to interoperability is validated and refined through a diverse set of pilot use cases, implemented across Spoke 2 and Spoke 3. These pilots serve as real-world testbeds for the deployed technical, semantic, organizational, and legal solutions, ensuring that the architecture and methodologies developed in Spoke 1 are robust, scalable, and responsive to the needs of end users.

The pilot studies encompass a broad spectrum of digital health innovation, including:

- **Digital twins and risk stratification tools:** BBCT workflow, a Digital Twin application for the estimating the risk of fracture of the femur and initially developed for the execution via a HPC scheduler (SLURM), was converted to be executed on a cloud infrastructure (EPIC). Virtual machines, with Rocky Linux 9 hardened operating system, were managed through OpenStack. Ansys was installed in a Docker container, and packages and libraries not needed by the particular execution of Mechanical APDL were removed: in this way the container size was reduced from 50 GB to 6 GB. Similarily, Python scripts were included in Docker containers specific for each microservice. BBCT workflow was then ported to Nextflow, configured to communicate with a Kubernetes cluster (RKE2 distribution, for enhanced security) composed by 3 master and 4 worker (64 GB RAM, 16 vCPU) nodes. A Network File System set up with Helm was used to exchange input and output files between

Commentato [SM14]: Vedere il mio commento in sezione 3.4, riguardo a riportare casi concreti di implementazione. Io posso inserire in coda una descrizione di come si stato fatto il mapping delle variabili in un paio di piloti. Ma ha senso inserirlo solo se ci sono altri casi applicativi concreti. Se non ci fossero, valuterei se non indicare solo che l'obiettivo è di validare il framework con i piloti. Rimandando al report successivo il risultato.

Commentato [BM15R14]: Io non ho le informazioni di dettaglio sui piloti, anche qui, bisognerebbe coinvolgere I loro responsabili per avere info piu' dettagliate. Vedi tu, se vuoi togliere. L'infrastruttura di interoperabilita' pero' c'e' ed e' a disposizione di tutti I piloti.

Commentato [SM16R14]: Ho sostituito le descrizioni con applicazioni verificate ed eventualmente difendibili.

Commentato [BM17]: @sabato verifica la correttezza per favore

Commentato [SM18R17]: Le categorie indicate sono coerenti

the nodes. Grafana and Prometheus were employed to monitor cluster activity. INFN licenses were used for Ansys execution during the test phase.

- **MLOps approaches in clinical applications:** the work involves the development and experimentation of a medical image analysis and generation pipeline based on MLOps practices, applied to the neuroimaging domain. In particular, a workflow was designed for the generation of synthetic 3D brain MRI images using Denoising Diffusion Probabilistic Models (DDPM), with the aim of mitigating the data scarcity typical of rare neurodegenerative diseases. The original images were acquired and managed according to clinical reference standards (DICOM), then converted and preprocessed into analysis formats (NIfTI), applying established procedures for normalization, registration to MNI space, and brain extraction. The model was trained and validated on publicly available multicenter datasets, with quantitative evaluation (MMD, FID, MS-SSIM with bootstrap confidence intervals) and qualitative assessment by neuroradiologists. The entire pipeline was implemented in a cloud/HPC environment, with attention to reproducibility, versioning, experiment traceability, and portability, laying the groundwork for future integration into regulated MLOps infrastructures for clinical applications.
- **Semi-automatic data mapping and standardization:** the project focuses on the development and evaluation of a pipeline leveraging Large Language Models (LLMs) for semi-automatic semantic standardization of clinical datasets from a custom format to HL7 FHIR, annotated with SNOMED CT and LOINC vocabularies, to improve interoperability and reduce manual effort. A workflow was devised in which structured source data are ingested, profiled, and converted into intermediate representations that capture field semantics and metadata, before being processed by LLM-driven mapping modules. Using techniques like embedding-based semantic retrieval and prompt engineering, the LLM suggests FHIR resource mappings and appropriate standardized codes for each data element, combining FHIR's structured resource model with internationally recognized clinical terminologies to ensure consistent representation. Human experts then review and refine these mappings to mitigate model errors and validate semantic accuracy, with feedback logged for reuse and iterative improvement. The entire system is implemented on the same cloud/HPC platform described in the reference, emphasizing reproducibility, version control, experiment traceability, and portability, and forming a foundation for future integration into regulated clinical data workflows that require HL7 FHIR compliance.
- .

By supporting these pilot studies, DARE generates critical feedback for continuous improvement.

Commentato [SM19]: Vedere commento precedente

Support of pilot use cases is performed making available the following platforms:
Multi-centric AI models validation platform: sensitive data from multiple research infrastructures is processed in compliance with GDPR and ethical requirements, enabling validation and retraining of AI models across federated sites. The platform is based on technologies like RKE2, Jupyter Hub, MLFlow.
Federated datalake for health research: EPIC Cloud enables the creation of a federated data lake, supporting scenarios such as central harvesting of remotely collected data, edge-level anonymization, federated learning, and secure feature extraction. This allows DARE to support diverse biomedical research communities with varying data governance needs. The datalake is based on technologies like RUCIO/FTS, XNAT and Keycloak.
EPIC Cloud is the backbone of a national datalake including all IRCCS for integrating omics and clinical data, supporting observational studies and large-scale analytics for disease prevention and personalized medicine. This platform is based on RedCap and on Molecular Tumor Board developd in the context of the Health Big Data project.
Medical imaging and genomics platform: EPIC Cloud and Salus Ratio integrates platforms such as XNAT for medical imaging, BOSCO for genomics [DOI: https://doi.org/10.22323/1.458.0042], Galaxy for bioinformatics workflows, and REDCap for electronic health records. These platforms are deployed with enhanced privacy and compliance controls, supporting DARE's commitment to ethical and legal requirements.

4.4. Lessons Learned and Best Practices

The DARE project's approach to interoperability is not merely theoretical—it is actively validated, stress-tested, and continuously refined through a diverse portfolio of pilot use cases implemented across Spoke 2 and Spoke 3. These pilots serve as real-world testbeds for the technical, semantic, organizational, and legal solutions developed in Spoke 1, ensuring that the project's architecture and methodologies are robust, scalable, and truly responsive to the needs of end users in the healthcare ecosystem.
A Living Laboratory for Digital Health Innovation: each pilot is designed to address concrete challenges in digital health, leveraging the interoperability framework to integrate data, processes, and stakeholders across organizational and technological boundaries. The pilots span a wide range of scenarios, from community-level prevention to advanced clinical decision support, and are selected to maximize both impact and the diversity of requirements. Pilots cover several aspects of prevention such as community-based

surveillance and prevention; digital twins and risk stratification tools; AI-driven analytics and decision support systems.

By supporting and closely monitoring these pilot studies, DARE not only demonstrates the practical value and adaptability of its interoperability framework but also generates a feedback loop for ongoing refinement and continuous improvement. Lessons learned from pilot implementation—such as integration bottlenecks, user experience challenges, and emerging data governance needs—are systematically captured and fed back into the project's development cycle. This iterative approach ensures that the solutions remain relevant, user-centered, and aligned with the evolving landscape of digital health.

The pilots also play a pivotal role in fostering stakeholder engagement and rigorous governance. Through participatory design, regular feedback sessions, and collaborative problem-solving, the pilots help align technical solutions with clinical workflows, organizational practices, and patient expectations. This dynamic, real-world validation process is essential for translating technical innovation into sustainable, impactful change across the digital health ecosystem.

5. Impact and Next Steps

5.1. Roadmap for Further Development

Building on the lessons learned from the initial deployment of interoperability solutions, the DARE project has outlined a clear and ambitious roadmap to ensure continued progress, scalability, and sustainability.

Extension of Interoperability Solutions to Additional Pilots and Partners the next phase will focus on broadening the reach of the established interoperability framework. This includes onboarding new pilot sites, research institutions, and healthcare partners, as well as integrating additional data sources and digital health platforms. By expanding the network of participants, DARE aims to foster a more comprehensive and representative ecosystem, further validating the robustness and adaptability of its solutions in diverse real-world contexts.

Continuous Alignment with Evolving National and European Standards recognizing the dynamic nature of the digital health landscape, DARE is committed to maintaining continuous alignment with the latest national and European standards, regulations, and best practices. This involves proactive monitoring of updates to frameworks such as HL7 FHIR, OMOP CDM, GDPR, EOSC and EHDS, as well as active participation in relevant

standardization bodies and policy forums. Regular reviews and iterative updates to the interoperability architecture will ensure ongoing compliance and future-proofing.

Ongoing Training and Capacity-Building for Stakeholder sustained impact requires that all stakeholders—clinicians, researchers, IT professionals, data stewards, and decision-makers—are equipped with the knowledge and skills to effectively use and contribute to interoperable systems. DARE will continue to invest in comprehensive training programs, workshops, and knowledge-sharing initiatives, fostering a culture of continuous learning and capacity-building across the ecosystem.

Preparation for the Next Deliverable (D3.3: Guidelines for Interoperability Governance) as part of its commitment to transparency and knowledge dissemination, DARE will synthesize the insights, methodologies, and best practices developed to date into the next key deliverable: D3.3 "Guidelines for Interoperability Governance." This document will serve as a reference for current and future partners, providing actionable guidance on governance models, stakeholder engagement, compliance strategies, and technical implementation.

Through these strategic actions, DARE is well-positioned to drive the ongoing evolution of digital health interoperability, ensuring that its solutions remain relevant, scalable, and impactful for years to come.

6. Conclusions

The deployment of multi-dimensional interoperability strategies and solutions within the DARE project, as detailed in this deliverable, has established a robust and future-oriented foundation for integrated, data-driven digital prevention in healthcare. By addressing the technical, semantic, organizational, legal, and ethical dimensions of interoperability, DARE will demostrate that it is possible to overcome the fragmentation of health information systems and enable seamless collaboration across a complex ecosystem of partners, platforms, and data sources.

The project's governance and coordination model has ensured strong alignment between enabling technologies and the real-world needs of pilot sites, fostering a culture of transparency, adaptability, and stakeholder engagement. The adoption of open standards, modular architectures, and federated cloud infrastructures—exemplified by the EPIC Cloud multisite deployment—has enabled secure, scalable, and compliant data integration and sharing. Semantic harmonization, achieved through the use of standardized terminologies and ontologies, has further empowered advanced analytics, AI-driven decision support, and secondary use of health data.

Organizational interoperability has been strengthened through the definition of shared workflows, the establishment of governance committees, and continuous capacity-building

initiatives. Legal and ethical compliance has been embedded at every level, leveraging regulatory sandboxes and robust consent management frameworks to ensure that innovation proceeds in a trustworthy and responsible manner.

Looking ahead, the roadmap for future development emphasizes the extension of interoperability solutions to new pilots and partners, continuous alignment with national and European standards, and ongoing training for all stakeholders. The forthcoming deliverable D3.3 will further consolidate the project's achievements by providing actionable guidelines for interoperability governance.

In summary, DARE's multi-dimensional approach to interoperability not only addresses current challenges but also positions the project as a reference model for sustainable, scalable, and ethically responsible digital health innovation at the national and European levels.