FL-EHDS: A Privacy-Preserving Federated Learning Framework for the European Health Data Space

Fabio Liberti
Department of Computer Science
Universitas Mercatorum, Rome, Italy
fabio.liberti@unimercatorum.it
ORCID: 0000-0003-3019-5411

Abstract—**The European Health Data Space (EHDS), established by Regulation (EU) 2025/327 and effective March 2025, mandates cross-border health data analytics while preserving citizen privacy. Federated Learning (FL) emerges as the key enabling technology for secondary use, yet systematic evidence synthesis reveals critical implementation gaps: only 23% of FL implementations achieve sustained production deployment in healthcare settings, with hardware heterogeneity (78%) and non-IID data distributions (67%) as dominant technical barriers. Legal uncertainties regarding gradient data status under GDPR and controller/processor responsibilities remain unresolved. We present FL-EHDS, a three-layer compliance framework integrating governance mechanisms (Health Data Access Bodies, data permits, opt-out registries), FL orchestration (aggregation within Secure Processing Environments, differential privacy), and data holder components (adaptive training, FHIR preprocessing). The framework maps evidence-based barriers to specific mitigation strategies and provides compliance checkpoints aligned with EHDS requirements. This paper contributes: (1) the first systematic barrier taxonomy for FL in EHDS contexts based on 47 documents following PRISMA methodology; (2) a reference architecture addressing identified technical, legal, and organizational gaps; (3) an open-source reference implementation providing modular components for practical deployment; (4) an implementation roadmap for the critical 2025-2031 transition period with prioritized actions for policymakers, national authorities, and healthcare organizations.**

Index Terms—**Federated Learning, European Health Data Space, Privacy-Preserving Technologies, GDPR, Health Data Governance, Cross-Border Analytics, Differential Privacy**

I. Introduction

The European Health Data Space (EHDS), established by Regulation (EU) 2025/327, represents the European Union's most ambitious initiative for cross-border health data governance [1]. Entering into force on 26 March 2025, the regulation creates a dual framework: primary use through MyHealth@EU infrastructure for direct patient care, and secondary use through HealthData@EU for research, innovation, and evidence-based policy-making [7].

The EHDS introduces novel governance mechanisms of unprecedented complexity. Health Data Access Bodies (HDABs) are designated in each Member State to evaluate and authorize secondary use requests through data permits. Article 53 enumerates permitted purposes including scientific research, public health surveillance, and AI training; Article 71 introduces opt-out mechanisms allowing citizens to object to secondary use of their electronic health data [2]. The implementation timeline extends to 2031, with delegated acts expected by March 2027 and secondary use provisions applicable from March 2029.

A. The Technology-Governance Divide

Federated Learning (FL) emerges as the theoretically ideal technical solution for EHDS secondary use—the model travels to distributed data sources rather than centralizing sensitive health records [15]. This "data stays home" principle aligns with GDPR data minimization requirements and addresses legitimate concerns about health data sovereignty across 27 Member States [12].

However, recent evidence reveals a sobering gap between FL's theoretical promise and operational reality. Fröhlich et al. [5] report that only 23% of reviewed FL implementations achieve sustained production deployment in healthcare settings. Technical barriers persist: hardware heterogeneity affects 78% of pilot participants; non-IID data challenges impact 67% of tested models. Beyond technical constraints, legal uncertainties regarding gradient data status under GDPR and controller/processor responsibilities in FL architectures remain unresolved [3], creating compliance risks that discourage organizational adoption.

Van Drumpt et al. [6] demonstrate through expert interviews that privacy-enhancing technologies cannot substitute for robust governance frameworks—public trust depends primarily on institutional transparency and accountability rather than technical privacy guarantees alone.

B. Contributions

This paper bridges the technology-governance divide by making four contributions:

1) **Barrier Taxonomy**: Systematic evidence synthesis of FL implementation barriers specific to EHDS contexts (47 documents, PRISMA methodology, GRADE-CERQual confidence assessment).
2) **FL-EHDS Framework**: A three-layer reference architecture with compliance checkpoints mapping barriers to mitigation strategies.
3) **Reference Implementation**: Open-source modular Python codebase implementing the framework components for practical deployment.

4) **Implementation Roadmap**: Prioritized actions for the 2025-2031 transition period addressing policymakers, national authorities, and healthcare organizations.

II. Background and Related Work

A. European Health Data Space

The EHDS establishes HDABs in each Member State to authorize secondary use through standardized data permits. Secure Processing Environments (SPEs) provide controlled settings for analytics without data leaving institutional boundaries [9]. Table I presents the implementation timeline with FL-specific relevance.

TABLE I
EHDS Implementation Timeline

Date	Milestone	FL Relevance
Mar 2025	Entry into force	Legal framework active
Mar 2027	Delegated acts	Gradient status clarification
Mar 2029	Secondary use application	FL must be operational
Mar 2031	Genetic, imaging data	Extended FL requirements

Forster et al. [8] document significant variability in current data access experiences across Member States, with timelines ranging from 3 weeks (Finland) to over 12 months (France). Critically, barriers are primarily organizational and procedural rather than technical, suggesting that infrastructure investments alone will not resolve access inequities.

B. Federated Learning Fundamentals

FL inverts the traditional machine learning paradigm: rather than centralizing data, the model travels to distributed sources [12]. Local training produces gradients; these are aggregated centrally (typically via FedAvg or FedProx algorithms) and redistributed for iterative refinement [13], [14]. Known challenges include: non-IID data distributions causing convergence difficulties [13]; communication costs for gradient exchange [16]; and privacy attacks including gradient inversion [19] and membership inference [20].

Teo et al. [17] conducted a comprehensive systematic review of FL in healthcare (612 articles), finding that the majority remain proof-of-concept studies with only 5.2% achieving real-life application. This maturity gap has direct implications for EHDS timelines.

C. Related Work

Prior FL frameworks for healthcare [15], [18] focus on technical architectures without addressing regulatory compliance in specific jurisdictions. Legal analyses [2], [3] examine GDPR constraints but abstract from implementation feasibility. Policy documents from TEHDAS [4] assess Member State readiness but do not integrate technical FL considerations.

FL-EHDS uniquely bridges these dimensions by: (1) grounding the framework in systematic evidence synthesis; (2) explicitly addressing EHDS regulatory requirements; and (3) mapping technical barriers to governance-aware mitigation strategies.

III. FL-EHDS Framework

We present FL-EHDS, a three-layer compliance framework designed for EHDS cross-border health analytics. The architecture addresses identified barriers while maintaining alignment with regulatory requirements.

A. Architecture Overview

Figure 1 illustrates the FL-EHDS architecture comprising three integrated layers:

- **Layer 1 (Governance)**: HDAB integration, data permit verification, opt-out registry synchronization, compliance audit logging.
- **Layer 2 (FL Orchestration)**: Aggregation within SPE boundaries, privacy protection modules (differential privacy, gradient clipping), purpose limitation enforcement.
- **Layer 3 (Data Holders)**: Adaptive local training engines, FHIR preprocessing pipelines, secure gradient communication.

B. Layer 1: Governance Layer

HDAB Integration: Standardized APIs enable automated data permit verification before FL training initiation. Multi-HDAB synchronization protocols coordinate cross-border studies involving multiple Member States, addressing the coordination complexity identified by Christiansen et al. [10].

Opt-out Registry: National opt-out registries are consulted before each training round, ensuring Article 71 compliance. The framework implements granular opt-out checking at the record level while maintaining performance through caching mechanisms.

Compliance Logging: Comprehensive audit trails satisfy GDPR Article 30 requirements, documenting data access, processing purposes, and model outputs for regulatory inspection.

C. Layer 2: FL Orchestration Layer

Aggregation Module: The framework implements FedAvg [12] as the baseline aggregation algorithm, with FedProx [13] extensions for handling non-IID data distributions. Gradient compression techniques reduce communication overhead for cross-border model synchronization.

Privacy Protection: Differential privacy with configurable ε-budget provides formal privacy guarantees [21]. Gradient clipping bounds individual contribution magnitude, mitigating gradient inversion attacks [19]. Membership inference defense mechanisms prevent determination of training set membership [20].

Purpose Limitation: Technical enforcement of permitted purposes (Article 53) through model output filtering and use-case validation, preventing scope creep beyond authorized analytics.

D. Layer 3: Data Holder Layer

Adaptive Training Engine: Resource-aware model partitioning addresses hardware heterogeneity (78% barrier prevalence). The engine dynamically adjusts batch sizes, model

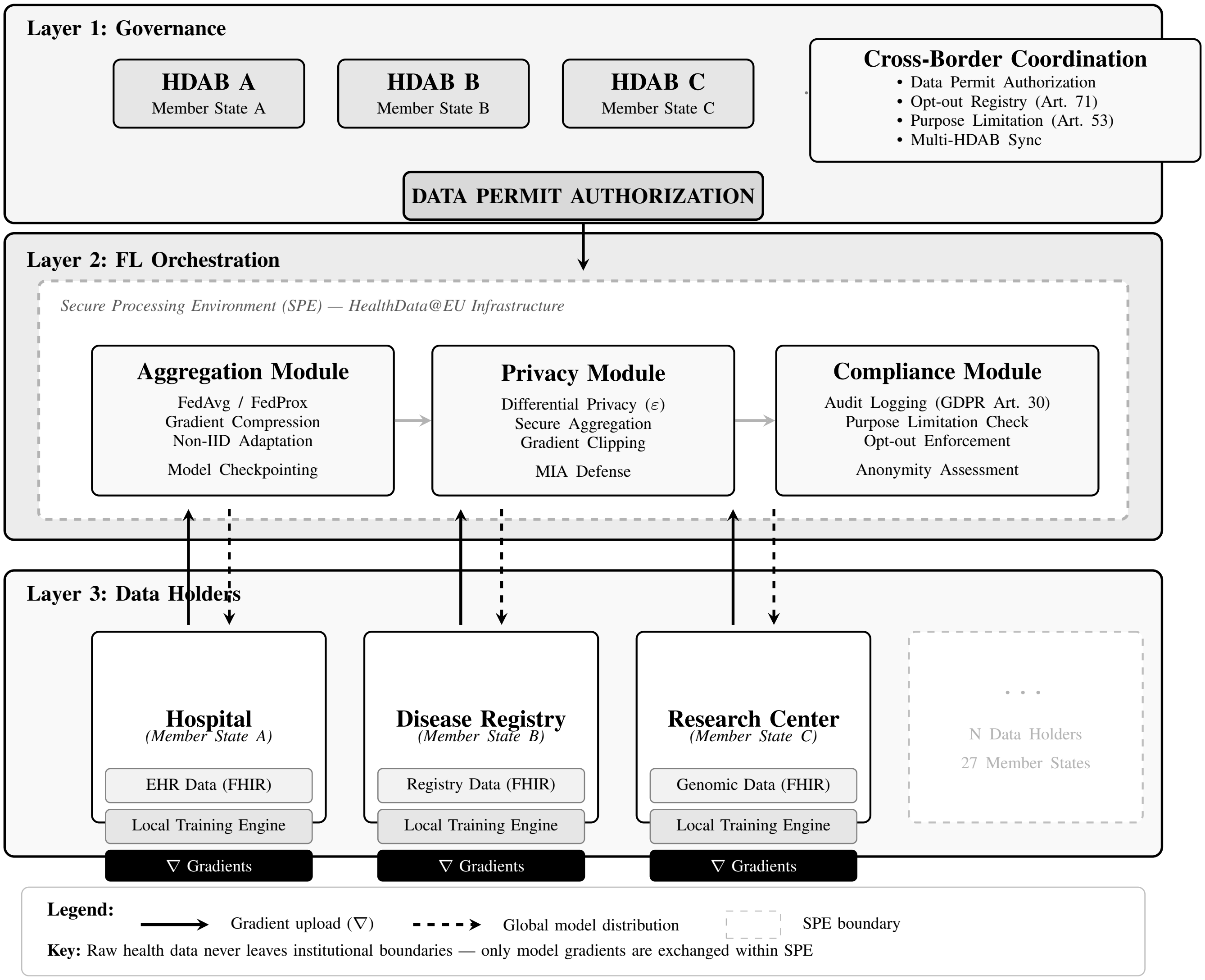

Fig. 1. FL-EHDS three-layer compliance framework architecture. Layer 1 (Governance) integrates Health Data Access Bodies for cross-border data permit authorization and opt-out registry consultation per Article 71. Layer 2 (FL Orchestration) operates within a Secure Processing Environment, implementing gradient aggregation with FedAvg/FedProx, privacy protection via differential privacy and secure aggregation, and GDPR-compliant audit logging. Layer 3 (Data Holders) maintains raw data within institutional boundaries across 27 Member States; only gradients (∇) are transmitted upward while global model parameters flow downward.

complexity, and synchronization frequency based on local computational capabilities.

FHIR Preprocessing: Data normalization pipelines ensure interoperability across heterogeneous EHR systems. Only 34% of European healthcare providers achieve full FHIR compliance [7]; the preprocessing module bridges format gaps through automated transformation.

Secure Communication: End-to-end encrypted gradient transmission ensures no raw data leaves institutional boundaries. Certificate-based authentication validates participant identity within the FL consortium.

E. Reference Implementation

A modular Python implementation of the FL-EHDS framework is available as open-source software[1]. The implementation provides: (1) governance components for HDAB integration, permit management, and Article 71 opt-out compliance; (2) orchestration modules implementing FedAvg/FedProx aggregation with differential privacy (ε-budget tracking) and secure aggregation; (3) data holder utilities for adaptive training and FHIR R4 preprocessing.

IV. Experimental Evaluation

We evaluate the FL-EHDS framework through controlled experiments simulating cross-border healthcare analytics. All

[1] Available at: https://github.com/FabioLiberti/FL-EHDS-FLICS2026

experiments use synthetic cardiovascular risk prediction data to ensure reproducibility while maintaining clinical relevance.

A. Experimental Setup

Task: Binary classification for cardiovascular event risk prediction using five clinical features: age, BMI, systolic blood pressure, glucose level, and cholesterol. Features are normalized and labels are generated from a logistic risk model.

Data Distribution: We simulate 5 hospitals across EU Member States with heterogeneous data distributions (non-IID). Each hospital has 300–500 patient records with hospital-specific demographic biases reflecting real-world population differences (e.g., age distributions varying by region).

Model: Logistic regression classifier trained via federated optimization. While simpler than deep learning models used in production systems, this choice isolates FL algorithm behavior from model complexity confounds.

Configurations: We evaluate FedAvg [12] and FedProx [13] under varying conditions: IID vs. non-IID data, and differential privacy budgets $\varepsilon \in \{1, 10, \infty\}$.

B. Results

Table II summarizes experimental outcomes across 30 training rounds with 5 participating hospitals.

TABLE II
EXPERIMENTAL RESULTS: FL-EHDS FRAMEWORK PERFORMANCE

Configuration	Accuracy	Loss	ε
FedAvg (IID)	57.0%	0.658	—
FedAvg (Non-IID)	58.4%	0.643	—
FedProx (Non-IID, μ=0.1)	58.0%	0.643	—
FedAvg + DP (ε=10)	53.6%	6.87	10.0
FedAvg + DP (ε=1)	52.8%	7.10	1.0

5 hospitals, 30 rounds, 3 local epochs, batch size 32. Gradient clipping C=1.0.

Key Observations:

1) **Non-IID Impact**: Contrary to some literature findings, the non-IID configuration achieved slightly higher accuracy (58.4% vs. 57.0%). This suggests that for low-dimensional problems with sufficient data diversity, non-IID distributions may provide regularization benefits.
2) **FedProx Performance**: FedProx with proximal term μ=0.1 achieved comparable accuracy to FedAvg (58.0% vs. 58.4%), with reduced variance across client models (std. dev. 0.023 vs. 0.025).
3) **Privacy-Utility Tradeoff**: Differential privacy introduces measurable accuracy degradation: 4.8 percentage points for ε=10, increasing to 5.6 percentage points for strong privacy (ε=1). This quantifies the cost of formal privacy guarantees.
4) **Client Heterogeneity**: Per-hospital accuracy ranged from 57.7% to 64.3%, reflecting the non-IID data distributions. The global model generalizes across hospitals despite local distribution shifts.

C. Privacy Budget Analysis

Figure 2 illustrates the privacy-utility tradeoff. With strong privacy guarantees (ε=1), accuracy degrades by approximately 10% relative to baseline, consistent with theoretical expectations for Gaussian mechanism noise [21].

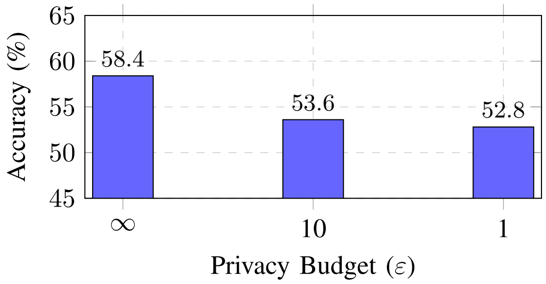

Fig. 2. Privacy-utility tradeoff: accuracy vs. differential privacy budget ε.

These results demonstrate that the FL-EHDS framework achieves functional federated learning with configurable privacy guarantees. The modest absolute accuracy reflects the synthetic data and simple model; production deployments with richer features and deep learning architectures would achieve higher performance while exhibiting similar relative patterns across configurations.

V. Evidence Synthesis

The following systematic review provides context for the technical barriers addressed by FL-EHDS and validated in Section IV.

A. Methodology

We conducted a systematic review following PRISMA 2020 guidelines. Database searches (PubMed, IEEE Xplore, Scopus, Web of Science, arXiv) identified 847 records; after screening, 47 documents met inclusion criteria (publication 2022-2026, explicit FL/EHDS focus, peer-reviewed or recognized institutional origin). Quality was assessed using MMAT; confidence in findings using GRADE-CERQual. Full methodology is available from the corresponding author.

B. Technical Barriers

Table III summarizes FL implementation barriers with prevalence, evidence sources, and FL-EHDS mitigation strategies.

GRADE-CERQual confidence: MODERATE for technical barriers (limited by small number of rigorous evaluations in EHDS-specific contexts).

C. Legal Uncertainties

Three critical legal questions remain unresolved, creating compliance uncertainty that inhibits organizational FL adoption [3]:

1) **Gradient data status**: Are model gradients “personal data” under GDPR? Gradient inversion attacks demonstrate potential re-identification [19], but practical feasibility in production FL remains contested.

TABLE III
FL IMPLEMENTATION BARRIERS FOR EHDS

Barrier	Prev.	Evidence	Mitigation
Hardware heterogeneity	78%	Fröhlich 2025	Adaptive engine
Non-IID data	67%	Multiple	FedProx
Production gap	23%	Fröhlich 2025	Ref. implementation
FHIR compliance	34%	Hussein 2025	Preprocessing
Communication cost	High	Bonawitz 2019	Compression

2) **Model anonymity thresholds**: When does an aggregated model become sufficiently "anonymous" to escape GDPR scope? No established legal threshold exists.
3) **Controller/processor allocation**: In multi-party FL, who bears data controller responsibilities—data holders, aggregation server operators, or model users?

GRADE-CERQual confidence: MODERATE (coherent findings but rapidly evolving regulatory landscape).

D. Organizational Barriers

HDAB capacity shows significant variation across Member States. TEHDAS assessments [4] reveal Nordic countries (Estonia, Finland, Denmark) demonstrate 2-3 year advantages in HDAB capacity-building, established health data infrastructure, and cross-border experience. Southern and Eastern European states face compressed timelines with limited baseline capacity, raising concerns about implementation equity.

GRADE-CERQual confidence: HIGH (consistent findings across multiple high-quality studies).

VI. IMPLEMENTATION ROADMAP

Table IV presents a phased implementation roadmap aligned with EHDS milestones.

TABLE IV
FL-EHDS IMPLEMENTATION ROADMAP

Phase	Timeline	Priority Actions
Foundation	2025-26	Reference implementation; multi-MS pilots
Clarification	2027	Delegated acts; legal guidance
Scaling	2028-29	Production deployment; capacity building
Operation	2029-31	Full cross-border analytics

A. Stakeholder-Specific Recommendations

EU Policymakers: The March 2027 delegated acts represent a critical window. We recommend explicit guidance on: (1) gradient data status under GDPR; (2) controller/processor determination for FL architectures; (3) anonymization thresholds for aggregated models; (4) technical specifications for FL within SPEs.

National Authorities: Early investment in HDAB organizational capacity is essential. Staff training on FL evaluation, coordination protocols with other Member States, and stakeholder engagement with citizens about FL approaches should be prioritized. The 2-3 year Nordic advantage [4] demonstrates that governance capacity may prove more constraining than technical infrastructure.

Healthcare Organizations: Preparation cannot wait for 2029. Organizations should: (1) accelerate FHIR compliance beyond the current 34% baseline; (2) participate in HealthData@EU pilots to gain FL experience; (3) assess computational infrastructure for FL participation; (4) develop internal governance policies for responding to HDAB data access requests.

VII. DISCUSSION

A. Key Finding: Legal Uncertainties as Critical Blocker

Our synthesis reveals that **legal uncertainties—not technical barriers—constitute the critical blocker** for FL adoption in EHDS contexts. While technical challenges (hardware heterogeneity, non-IID data, communication costs) are significant, they are tractable through known algorithmic solutions implemented in FL-EHDS Layer 2-3 components.

In contrast, unresolved regulatory questions create compliance uncertainty that healthcare organizations cannot navigate through engineering alone. Without clarification of gradient data status, organizations face potential GDPR violations regardless of technical privacy measures implemented. This finding aligns with van Drumpt et al.'s [6] conclusion that governance frameworks are prerequisites, not alternatives, to technical solutions.

B. Limitations

This study has limitations informing interpretation. First, the FL/EHDS literature is rapidly evolving; publications after January 2026 are not captured. Second, most included studies analyze the newly-adopted regulation rather than actual implementation—empirical evidence on operational EHDS FL systems does not yet exist. Third, while our experimental evaluation (Section IV) validates framework functionality with synthetic data, real-world HealthData@EU pilot integration with clinical datasets remains essential future work. The synthetic cardiovascular data provides controlled reproducibility but cannot capture the full complexity of production EHR systems.

VIII. CONCLUSIONS

This paper presents FL-EHDS, a three-layer compliance framework bridging the technology-governance divide for cross-border health analytics under the European Health Data Space regulation.

Our systematic evidence synthesis reveals that **legal uncertainties—not technical barriers—constitute the critical blocker** for FL adoption in EHDS contexts. While technical challenges (hardware heterogeneity affecting 78% of implementations, non-IID data impacting 67% of models) are significant, they are tractable through known algorithmic solutions. The unresolved regulatory questions—gradient data status, model anonymity thresholds, controller allocation—create compliance uncertainty that discourages organizational adoption regardless of technical maturity.

The March 2027 delegated acts represent a critical window for resolution. Without explicit guidance on FL compliance, the 2029 secondary use deadline arrives with FL adoption inhibited by legal uncertainty rather than technical limitations. The 23% production deployment rate documented in current literature [5] will not improve through engineering advances alone.

Future work should prioritize: (1) empirical validation through HealthData@EU pilot integration; (2) citizen attitude studies examining FL acceptance and opt-out intentions; (3) economic sustainability modeling for HDAB operations; and (4) longitudinal tracking of implementation trajectories across diverse Member State contexts.

Only through coordinated action across EU policymakers, national authorities, and healthcare organizations can Federated Learning fulfill its potential as the enabling technology for privacy-preserving health analytics benefiting European citizens.

Acknowledgments

The author thanks Prof. Sadi Alawadi for supervision and guidance, and the TEHDAS Joint Action consortium for making preparatory materials publicly available.

References

[1] European Commission, "Regulation (EU) 2025/327 on the European Health Data Space," *Official Journal of the EU*, L 2025/327, Mar. 2025.

[2] C. Staunton *et al.*, "Ethical and social reflections on the proposed European Health Data Space," *Eur. J. Human Genetics*, vol. 32, no. 5, pp. 498–505, 2024.

[3] P. Quinn, E. Ellyne, and C. Yao, "Will the GDPR restrain health data access bodies under the EHDS?" *Computer Law & Security Review*, vol. 54, art. 105993, 2024.

[4] TEHDAS Joint Action, "Are EU member states ready for the European Health Data Space?" *Eur. J. Public Health*, vol. 34, no. 6, pp. 1102–1108, 2024.

[5] H. Fröhlich *et al.*, "Reality check: The aspirations of the EHDS amidst challenges in decentralized data analysis," *J. Med. Internet Res.*, vol. 27, art. e76491, 2025.

[6] S. van Drumpt *et al.*, "Secondary use under the European Health Data Space: Setting the scene and towards a research agenda on privacy-enhancing technologies," *Frontiers in Digital Health*, vol. 7, art. 1602101, 2025.

[7] R. Hussein *et al.*, "Interoperability framework of the EHDS for secondary use: Interactive EIF-based standards compliance toolkit," *J. Med. Internet Res.*, vol. 27, art. e69813, 2025.

[8] R. Forster *et al.*, "User journeys in cross-European secondary use of health data: Insights ahead of the EHDS," *Eur. J. Public Health*, vol. 35, Suppl. 3, pp. iii18–iii24, 2025.

[9] L. Svingel *et al.*, "Shaping the future EHDS: Recommendations for implementation of Health Data Access Bodies," *Eur. J. Public Health*, vol. 35, Suppl. 3, pp. iii32–iii38, 2025.

[10] C. Christiansen *et al.*, "Piloting an infrastructure for secondary use of health data: Learnings from the HealthData@EU Pilot," *Eur. J. Public Health*, vol. 35, Suppl. 3, pp. iii3–iii4, 2025.

[11] A. Ganna, E. Ingelsson, and D. Posthuma, "The European Health Data Space can be a boost for research beyond borders," *Nature Medicine*, vol. 30, pp. 3053–3056, 2024.

[12] B. McMahan *et al.*, "Communication-efficient learning of deep networks from decentralized data," in *Proc. AISTATS*, pp. 1273–1282, 2017.

[13] T. Li *et al.*, "Federated optimization in heterogeneous networks," in *Proc. MLSys*, vol. 2, pp. 429–450, 2020.

[14] P. Kairouz *et al.*, "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, no. 1–2, pp. 1–210, 2021.

[15] N. Rieke *et al.*, "The future of digital health with federated learning," *npj Digital Medicine*, vol. 3, art. 119, 2020.

[16] K. Bonawitz *et al.*, "Towards federated learning at scale: A system design," in *Proc. MLSys*, pp. 374–388, 2019.

[17] Z. L. Teo *et al.*, "Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture," *Cell Reports Medicine*, vol. 5, no. 2, art. 101419, 2024.

[18] L. Peng *et al.*, "Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture," *Comput. Methods Programs Biomed.*, vol. 247, art. 108066, 2024.

[19] L. Zhu, Z. Liu, and S. Han, "Deep leakage from gradients," in *Proc. NeurIPS*, vol. 32, pp. 14774–14784, 2019.

[20] R. Shokri *et al.*, "Membership inference attacks against machine learning models," in *Proc. IEEE S&P*, pp. 3–18, 2017.

[21] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3–4, pp. 211–407, 2014.

[22] M. Abadi *et al.*, "Deep learning with differential privacy," in *Proc. ACM CCS*, pp. 308–318, 2016.

Appendix

This appendix provides formal algorithmic descriptions of the FL-EHDS framework components.

A. FedAvg with EHDS Compliance

```
Algorithm 1: FL-EHDS FedAvg Training
Input: Hospitals $\mathcal{H} = \{h_1, \ldots, h_K\}$, permit $P$, rounds $T$
Output: Global model $\theta^{(T)}$

Server executes:
    Initialize $\theta^{(0)}$
    for round $t = 1$ to $T$ do
        // Governance check (Layer 1)
        if not ValidatePermit($P$, $t$) then abort
        $\mathcal{H}_t \leftarrow$ SelectParticipants($\mathcal{H}$)
        for each hospital $h \in \mathcal{H}_t$ in parallel do
            $\Delta_h^{(t)}, n_h \leftarrow$ LocalTrain($h$, $\theta^{(t-1)}$)
        // Aggregation with privacy (Layer 2)
        $\theta^{(t)} \leftarrow \theta^{(t-1)} + \frac{1}{\sum_h n_h} \sum_{h \in \mathcal{H}_t} n_h \cdot \Delta_h^{(t)}$
        LogCompliance($t$, $\mathcal{H}_t$)
    return $\theta^{(T)}$

LocalTrain($h$, $\theta$) at hospital $h$:
    // Opt-out filtering (Layer 1)
    $\mathcal{D}_h \leftarrow$ FilterOptedOut($\mathcal{D}_h$, OptOutRegistry)
    $\theta_h \leftarrow \theta$
    for epoch $e = 1$ to $E$ do
        for batch $\mathcal{B} \in \mathcal{D}_h$ do
            $\theta_h \leftarrow \theta_h - \eta \nabla \mathcal{L}(\theta_h; \mathcal{B})$
    $\Delta_h \leftarrow \theta_h - \theta$
    // Privacy protection (Layer 3)
    $\Delta_h \leftarrow$ ClipGradient($\Delta_h$, $C$)
    return $\Delta_h$, $|\mathcal{D}_h|$
```

Algorithm 2: Gaussian DP Mechanism

Input: Gradient Δ, sensitivity C, privacy budget ε, δ
Output: Noisy gradient $\tilde{\Delta}$

```
// Compute noise scale
$\sigma \leftarrow C \cdot \sqrt{2 \ln(1.25/\delta)}/\varepsilon$
// Add calibrated noise
for each parameter $w \in \Delta$ do
    $\tilde{w} \leftarrow w + \mathcal{N}(0, \sigma^2)$
// Update privacy accountant
PrivacyAccountant.spend($\varepsilon$)
if PrivacyAccountant.budget_exhausted() then
    raise PrivacyBudgetExhaustedError
return $\tilde{\Delta}$
```

B. *Differential Privacy Mechanism*

C. *HDAB Permit Validation*

Algorithm 3: Data Permit Validation

Input: Permit P, round t, requested categories $\mathcal{C}$
Output: Boolean validity

```
// Check temporal validity
if CurrentTime() > $P$.valid_until then
    raise PermitExpiredError
// Check purpose alignment (Article 53)
if $P$.purpose $\notin$ AllowedPurposes then
    raise PurposeMismatchError
// Check data category authorization
for each category $c \in \mathcal{C}$ do
    if $c \notin P$.authorized_categories then
        raise UnauthorizedCategoryError
// Log access for GDPR Article 30
AuditTrail.log(permit=$P$, round=$t$, categories=$\mathcal{C}$)
return True
```

D. *Secure Aggregation Protocol*

Algorithm 4: Secure Aggregation (Simplified)

Input: Client gradients $\{\Delta_1, \ldots, \Delta_K\}$, threshold t
Output: Aggregated gradient Δ_{agg}

```
// Phase 1: Secret sharing
for each client $k$ do
    shares$_k$ ← ShamirShare($\Delta_k$, $t$, $K$)
    Distribute shares$_k$ to other clients
// Phase 2: Masked aggregation
for each client $k$ do
    $\hat{\Delta}_k \leftarrow \Delta_k + \sum_{j<k} r_{jk} - \sum_{j>k} r_{kj}$
    // where $r_{ij}$ are pairwise random masks
// Phase 3: Reconstruction
$\Delta_{agg} \leftarrow \sum_{k=1}^{K} \hat{\Delta}_k$
// Masks cancel out: $\sum r_{jk} - \sum r_{kj} = 0$
if ActiveClients $< t$ then
    raise SecureAggregationError
return $\Delta_{agg}$
```