COMPUTER SCIENCE BULLETIN

VOL: 09 ISSUE: 01 (2026)

E-ISSN: 2959-5347 P-ISSN: 2959-5339

Privacy-Preserving Federated Learning for Secure Healthcare Data Analysis in Distributed Computing Environments

Andreas Holzinger

University of Natural Resources and Life Sciences, Vienna, Austria

Abstract. *Healthcare analytics increasingly depends on geographically distributed clinical records, connected medical devices, imaging repositories, and institutional data platforms. Conventional centralized learning can increase exposure of sensitive information and amplify governance challenges when data must cross organizational boundaries. Federated learning offers a collaborative alternative in which participating sites train local models and exchange updates rather than raw records. However, model updates can still disclose private information, and operational heterogeneity can compromise accuracy, fairness, and reliability. This narrative review develops a privacy-preserving federated learning framework for secure healthcare data analysis, combining local preprocessing, identity and consent controls, secure aggregation, differential privacy, authenticated communication, robust aggregation, and continuous audit. It evaluates design trade-offs across clinical utility, confidentiality, interoperability, deployment cost, and reproducibility. The proposed framework is conceptual rather than an experimentally validated implementation; its evaluation criteria and illustrative figure are intended to guide future empirical studies. A governance-centered design is recommended in which technical protections complement clinical oversight and institutional accountability.*

Keywords: *federated learning, healthcare privacy, secure aggregation, differential privacy, distributed computing*

Introduction

Digital healthcare generates highly heterogeneous information, including electronic health records, medical images, laboratory observations, prescriptions, wearable sensor streams, and clinical narratives. These sources can support early diagnosis, risk prediction, operational planning, and population health research. Yet hospitals and laboratories operate under distinct legal obligations, technical standards, and patient expectations, making unrestricted pooling of identifiable records difficult to justify.Federated learning (FL) distributes model optimization across data holders and keeps the underlying patient records at their source. In a typical round, a coordinating service distributes a model, participating institutions train it locally, and a protected aggregation service combines updates. The arrangement reduces routine data movement but does not by itself establish privacy, security, or equitable model performance. Gradient leakage,

membership inference, poisoned updates, and uneven patient populations remain serious concerns. This article therefore examines FL as a socio-technical system rather than treating decentralization as an automatic guarantee of confidentiality.

Healthcare Data Sensitivity and Distributed Analytics:

Clinical information is unusually sensitive because diagnostic codes, imaging, longitudinal encounters, genomic profiles, and treatment histories can identify or stigmatize patients even when obvious identifiers are removed.Clinical information is unusually sensitive because diagnostic codes, imaging, longitudinal encounters, genomic profiles, and treatment histories can identify or stigmatize patients even when obvious identifiers are removed. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Data minimization should start with a clear clinical question, an approved feature set,documented retention limits, and a lawful basis for processing. Hospitals should distinguish direct care, service improvement, and secondary research because consent and authorization requirements differ.Data minimization should start with a clear clinical question, an approved feature set, documented retention limits, and a lawful basis for processing. Hospitals should distinguish direct care, service improvement, and secondary research because consent and authorization requirements differ. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Distributed analytics can reduce transfers between hospitals, but differences in coding systems, acquisition equipment, language, and clinical workflow create hidden inconsistencies. Common data dictionaries, provenance records, and documented exclusion rules help ensure that model updates represent comparable phenomena.

Healthcare Data Sensitivity and Distributed Analytics:

Clinical information is unusually sensitive because diagnostic codes, imaging, longitudinal encounters, genomic profiles, and treatment histories can identify or stigmatize patients even when obvious identifiers are removed. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Clinical information is unusually sensitive because diagnostic codes, imaging, longitudinal encounters, genomic profiles, and treatment histories can identify or stigmatize patients even when obvious identifiers are removed. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Data minimization should start with a clear clinical question, an approved feature set, documented retention limits, and a lawful basis for processing. Hospitals should distinguish direct care, service improvement, and secondary research because consent and authorization requirements differ. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Data minimization should start with a clear clinical question, an approved feature set, documented retention limits, and a lawful basis for processing. Hospitals should distinguish direct care, service improvement, and secondary research because consent and authorization requirements differ. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Distributed analytics can reduce transfers between hospitals, but differences in coding systems, acquisition equipment, language, and clinical workflow create hidden inconsistencies. Common data dictionaries, provenance records, and documented exclusion rules

help ensure that model updates represent comparable phenomena. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Distributed analytics can reduce transfers between hospitals, but differences in coding systems, acquisition equipment, language, and clinical workflow create hidden inconsistencies. Common data dictionaries, provenance records, and documented exclusion rules help ensure that model updates represent comparable phenomena. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.A realistic architecture also accounts for local data stewards, patient rights, access requests, incident response, and audit trails. Governance should specify who may initiate training, inspect metrics, deploy models, or withdraw an institution from collaboration. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: A realistic architecture also accounts for local data stewards, patient rights, access requests, incident response, and audit trails. Governance should specify who may initiate training, inspect metrics, deploy models, or withdraw an institution from collaboration. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.

Federated Learning Architecture and Workflow:

A cross-silo federation connects a manageable set of hospitals or research institutions, whereas cross-device FL may involve large numbers of mobile or wearable devices with intermittent connectivity.A cross-silo federation connects a manageable set of hospitals or research institutions, whereas cross-device FL may involve large numbers of mobile or wearable devices with intermittent connectivity. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.The coordinator distributes a versioned global model and an authenticated training configuration. Each institution checks eligibility, preprocesses records, trains on local partitions, and returns protected updates and approved summary metrics.The coordinator distributes a versioned global model and an authenticated training configuration. Each institution checks eligibility, preprocesses records, trains on local partitions, and returns protected updates and approved summary metrics. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.FedAvg combines client weights in proportion to local sample counts, but this approach can favor large hospitals and can become unstable when local distributions are strongly non-identical. Alternatives include proximal regularization, personalized heads, clustered training, and adaptive server optimization.FedAvg combines client weights in proportion to local sample counts, but this approach can favor large hospitals and can become unstable when local distributions are strongly non-identical. Alternatives include proximal regularization, personalized heads, clustered training, and adaptive server optimization. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Operational design must address failed rounds, model rollback, incompatible software, bandwidth constraints, and reproducible random seeds. Signed artifacts, model registries, and explicit version compatibility checks reduce accidental or malicious inconsistency.Operational design must address failed rounds, model rollback, incompatible

software, bandwidth constraints, and reproducible random seeds. Signed artifacts, model registries, and explicit version compatibility checks reduce accidental or malicious inconsistency. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.

Federated Learning Architecture and Workflow:

A cross-silo federation connects a manageable set of hospitals or research institutions, whereas cross-device FL may involve large numbers of mobile or wearable devices with intermittent connectivity. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: A cross-silo federation connects a manageable set of hospitals or research institutions, whereas cross-device FL may involve large numbers of mobile or wearable devices with intermittent connectivity. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.The coordinator distributes a versioned global model and an authenticated training configuration. Each institution checks eligibility, preprocesses records, trains on local partitions, and returns protected updates and approved summary metrics. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: The coordinator distributes a versioned global model and an authenticated training configuration. Each institution checks eligibility, preprocesses records, trains on local partitions, and returns protected updates and approved summary metrics. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.FedAvg combines client weights in proportion to local sample counts, but this approach can favor large hospitals and can become unstable when local distributions are strongly non-identical. Alternatives include proximal regularization, personalized heads, clustered training, and adaptive server optimization. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: FedAvg combines client weights in proportion to local sample counts, but this approach can favor large hospitals and can become unstable when local distributions are strongly non-identical. Alternatives include proximal regularization, personalized heads, clustered training, and adaptive server optimization. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Operational design must address failed rounds, model rollback, incompatible software, bandwidth constraints, and reproducible random seeds. Signed artifacts, model registries, and explicit version compatibility checks reduce accidental or malicious inconsistency. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.

Privacy Threats and Adversarial Models:

Keeping raw records local does not prevent all disclosure. Model inversion, membership inference, gradient reconstruction, and property inference can exploit information embedded in shared parameters or updates.Keeping raw records local does not prevent all disclosure. Model inversion, membership inference, gradient reconstruction, and property inference can exploit information embedded in shared parameters or updates. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance

criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Threat modeling should distinguish honest-but-curious coordinators, malicious participating sites, external network attackers, compromised endpoints, and colluding participants. The trust assumptions behind each protection should be stated explicitly.Threat modeling should distinguish honest-but-curious coordinators, malicious participating sites, external network attackers, compromised endpoints, and colluding participants. The trust assumptions behind each protection should be stated explicitly. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.An adversary may infer whether a rare diagnosis occurs at a small institution, reconstruct characteristics of a training example, or inject malicious updates to bias clinical predictions. Small client cohorts and repeated participation can make privacy attacks more effective.An adversary may infer whether a rare diagnosis occurs at a small institution, reconstruct characteristics of a training example, or inject malicious updates to bias clinical predictions. Small client cohorts and repeated participation can make privacy attacks more effective. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Mitigations should be layered: transport encryption protects messages in transit, secure aggregation limits visibility of individual updates, differential privacy bounds contribution leakage, and anomaly detection helps identify poisoned training behavior.Mitigations should be layered: transport encryption protects messages in transit, secure aggregation limits visibility of individual updates, differential privacy bounds contribution leakage, and anomaly detection helps identify poisoned training behavior. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.

Secure Aggregation and Cryptographic Protection:

Secure aggregation is designed to reveal an aggregate update while hiding individual client contributions from the server under a specified adversarial and dropout model. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Secure aggregation is designed to reveal an aggregate update while hiding individual client contributions from the server under a specified adversarial and dropout model. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Protocols commonly use masking and cryptographic key exchange so that masks cancel in the sum. Practical deployments must account for participant dropout, key rotation, authentication, replay resistance, and the risk of collusion. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Protocols commonly use masking and cryptographic key exchange so that masks cancel in the sum. Practical deployments must account for participant dropout, key rotation, authentication, replay resistance, and the risk of collusion. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Homomorphic encryption can enable selected computations on ciphertext but introduces computational and communication costs. Trusted execution environments offer

another option, though they depend on hardware trust and protection against implementation vulnerabilities. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Homomorphic encryption can enable selected computations on ciphertext but introduces computational and communication costs. Trusted execution environments offer another option, though they depend on hardware trust and protection against implementation vulnerabilities. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Cryptographic safeguards should be evaluated against explicit attack assumptions rather than advertised as universal guarantees. Key management, recovery procedures, and independent security reviews are as important as selecting a particular protocol. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Cryptographic safeguards should be evaluated against explicit attack assumptions rather than advertised as universal guarantees. Key management, recovery procedures, and independent security reviews are as important as selecting a particular protocol. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.

Differential Privacy and Utility Trade-offs:

Differential privacy provides a mathematical framework for bounding the influence of an individual record on released information, typically through calibrated random noise and a declared privacy budget.Differential privacy provides a mathematical framework for bounding the influence of an individual record on released information, typically through calibrated random noise and a declared privacy budget. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.In federated optimization, client-level or record-level privacy can be pursued through clipping and noisy aggregation. These settings protect different units and should not be conflated in reports to clinicians or regulators.In federated optimization, client-level or record-level privacy can be pursued through clipping and noisy aggregation. These settings protect different units and should not be conflated in reports to clinicians or regulators. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Stronger privacy can reduce sensitivity to individual examples but may lower predictive performance, particularly for minority classes, rare diseases, and small institutions. A meaningful evaluation should report privacy parameters alongside utility and subgroup performance.Stronger privacy can reduce sensitivity to individual examples but may lower predictive performance, particularly for minority classes, rare diseases, and small institutions. A meaningful evaluation should report privacy parameters alongside utility and subgroup performance. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Privacy accounting must consider repeated rounds, adaptive model selection, hyperparameter tuning, and public release of metrics. Teams should not claim formal privacy without identifying the exact mechanism, neighboring-data definition, and accounting method.Privacy accounting must consider repeated rounds, adaptive model selection, hyperparameter tuning, and public release of metrics. Teams should not claim formal privacy

without identifying the exact mechanism, neighboring-data definition, and accounting method. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.

Differential Privacy and Utility Trade-offs:

Differential privacy provides a mathematical framework for bounding the influence of an individual record on released information, typically through calibrated random noise and a declared privacy budget. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Differential privacy provides a mathematical framework for bounding the influence of an individual record on released information, typically through calibrated random noise and a declared privacy budget. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.In federated optimization, client-level or record-level privacy can be pursued through clipping and noisy aggregation. These settings protect different units and should not be conflated in reports to clinicians or regulators. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: In federated optimization, client-level or record-level privacy can be pursued through clipping and noisy aggregation. These settings protect different units and should not be conflated in reports to clinicians or regulators. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Stronger privacy can reduce sensitivity to individual examples but may lower predictive performance, particularly for minority classes, rare diseases, and small institutions. A meaningful evaluation should report privacy parameters alongside utility and subgroup performance. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Stronger privacy can reduce sensitivity to individual examples but may lower predictive performance, particularly for minority classes, rare diseases, and small institutions. A meaningful evaluation should report privacy parameters alongside utility and subgroup performance. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Privacy accounting must consider repeated rounds, adaptive model selection, hyperparameter tuning, and public release of metrics. Teams should not claim formal privacy without identifying the exact mechanism, neighboring-data definition, and accounting method. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Privacy accounting must consider repeated rounds, adaptive model selection, hyperparameter tuning, and public release of metrics. Teams should not claim formal privacy without identifying the exact mechanism, neighboring-data definition, and accounting method. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.

Non-IID Data, Fairness, and Model Robustness:

Hospitals differ in case mix, referral pathways, socioeconomic conditions, diagnostic practices, scanner vendors, and record completeness. Such non-identically distributed data can cause a global model to underperform at specific sites.Hospitals differ in case mix, referral pathways,

socioeconomic conditions, diagnostic practices, scanner vendors, and record completeness. Such non-identically distributed data can cause a global model to underperform at specific sites. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Evaluation should report per-institution and subgroup metrics, calibration, false-positive and false-negative rates, and uncertainty rather than relying only on pooled accuracy. External validation is essential when the deployment population differs from training participants.Evaluation should report per-institution and subgroup metrics, calibration, false-positive and false-negative rates, and uncertainty rather than relying only on pooled accuracy. External validation is essential when the deployment population differs from training participants. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Robust aggregation, client weighting constraints, and local personalization may help mitigate outliers, but aggressive rejection of unusual updates can also suppress legitimate minority populations. Statistical outliers are not automatically adversarial.Robust aggregation, client weighting constraints, and local personalization may help mitigate outliers, but aggressive rejection of unusual updates can also suppress legitimate minority populations. Statistical outliers are not automatically adversarial. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.

Non-IID Data, Fairness, and Model Robustness:

Hospitals differ in case mix, referral pathways, socioeconomic conditions, diagnostic practices, scanner vendors, and record completeness. Such non-identically distributed data can cause a global model to underperform at specific sites. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Hospitals differ in case mix, referral pathways, socioeconomic conditions, diagnostic practices, scanner vendors, and record completeness. Such non-identically distributed data can cause a global model to underperform at specific sites. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Evaluation should report per-institution and subgroup metrics, calibration, false-positive and false-negative rates, and uncertainty rather than relying only on pooled accuracy. External validation is essential when the deployment population differs from training participants. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Evaluation should report per-institution and subgroup metrics, calibration, false-positive and false-negative rates, and uncertainty rather than relying only on pooled accuracy. External validation is essential when the deployment population differs from training participants. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.Robust aggregation, client weighting constraints, and local personalization may help mitigate outliers, but aggressive rejection of unusual updates can also suppress legitimate minority populations. Statistical outliers are not automatically adversarial. This issue becomes especially consequential when a federation grows beyond a small pilot and includes institutions with different resource levels and regulatory obligations.Implementation implication: Robust

aggregation, client weighting constraints, and local personalization may help mitigate outliers, but aggressive rejection of unusual updates can also suppress legitimate minority populations. Statistical outliers are not automatically adversarial. Teams should test the assumption under realistic failure conditions, record the outcome in a reproducible audit log, and establish a process for revisiting the choice after deployment.

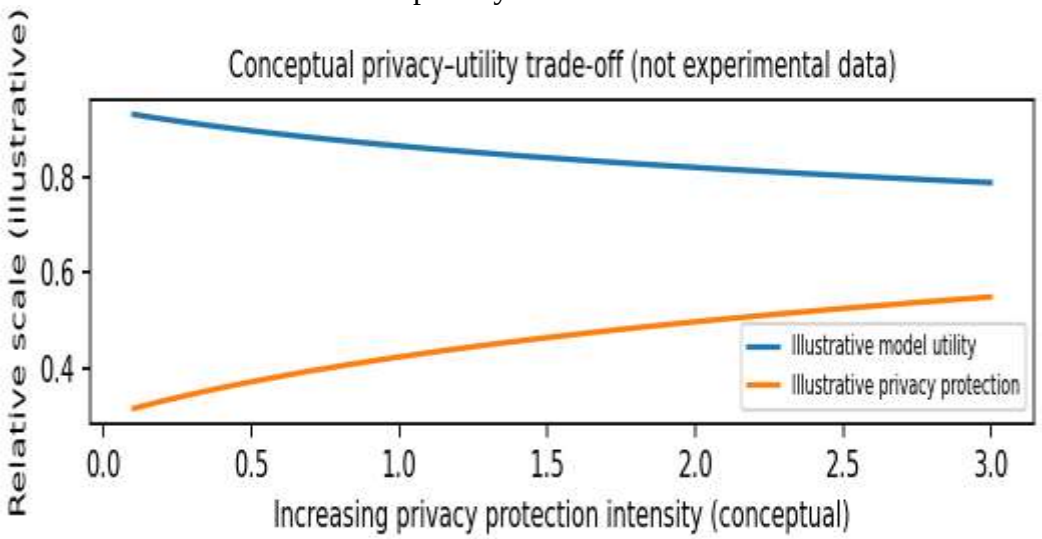
Interoperability, Infrastructure, and Deployment:

Electronic health record systems differ in data models, interfaces, terminology, and access control. Federated training requires semantic consistency as well as secure transport.Electronic health record systems differ in data models, interfaces, terminology, and access control. Federated training requires semantic consistency as well as secure transport. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Standards-based mappings, such as appropriately implemented FHIR resources and standardized clinical vocabularies, can support data harmonization, although local validation remains necessary for each mapping.Standards-based mappings, such as appropriately implemented FHIR resources and standardized clinical vocabularies, can support data harmonization, although local validation remains necessary for each mapping. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Deployment planning should include network reliability, accelerators, secure enclaves where appropriate, container isolation, credential rotation, logging, storage quotas, and cost attribution. Smaller institutions may require lightweight models and asynchronous participation.Deployment planning should include network reliability, accelerators, secure enclaves where appropriate, container isolation, credential rotation, logging, storage quotas, and cost attribution. Smaller institutions may require lightweight models and asynchronous participation. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Institutional review boards, security officers, clinical owners, and technical operators should approve a shared runbook defining onboarding, training authorization, audit retention, breach response, and withdrawal procedures.Institutional review boards, security officers, clinical owners, and technical operators should approve a shared runbook defining onboarding, training authorization, audit retention, breach response, and withdrawal procedures. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.

Evaluation Methodology and Reporting:

A credible empirical evaluation should compare federated models with site-only baselines and a centralized reference model when a lawful and ethically approved comparison is feasible.A credible empirical evaluation should compare federated models with site-only baselines and a centralized reference model when a lawful and ethically approved comparison is feasible. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Use patient-level separation between training and testing, temporally separated validation where possible, and external test hospitals. Avoid leakage through repeated encounters,

correlated family records, and site-specific artifacts.Use patient-level separation between training and testing, temporally separated validation where possible, and external test hospitals. Avoid leakage through repeated encounters, correlated family records, and site-specific artifacts. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Report discrimination, calibration, sensitivity, specificity, subgroup equity, privacy budget, update size, training time, energy use, client dropout, and attack resilience. Confidence intervals and prespecified endpoints improve interpretability.Report discrimination, calibration, sensitivity, specificity, subgroup equity, privacy budget, update size, training time, energy use, client dropout, and attack resilience. Confidence intervals and prespecified endpoints improve interpretability. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.Because this article does not present a collected dataset or executed training experiment, no accuracy, latency, or privacy guarantee is claimed. The accompanying graph is a conceptual illustration of a hypothesized trade-off, not measured evidence.Because this article does not present a collected dataset or executed training experiment, no accuracy, latency, or privacy guarantee is claimed. The accompanying graph is a conceptual illustration of a hypothesized trade-off, not measured evidence. In practice, this consideration should be translated into a documented decision rule, a responsible owner, and a measurable acceptance criterion. Such documentation allows participating institutions to identify differences before they become silent sources of model error or privacy risk.



Summary:

Privacy-preserving federated learning can support collaborative healthcare analytics without routinely transferring source records, but the locality of data alone does not ensure confidentiality, security, clinical validity, or fairness. A defensible system integrates clear governance, authenticated training, secure aggregation, appropriately accounted differential privacy, robust optimization, subgroup evaluation, and continuous monitoring. Design choices must be evaluated together: stronger privacy may alter predictive utility, complex cryptography may affect participation, and global accuracy may conceal poor outcomes for underrepresented

patients. The framework in this narrative review provides a structured basis for empirical validation, not a claim of achieved clinical performance.

References:

- McMahan, H. B., et al. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS.
- Kairouz, P., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1–2), 1–210.
- Rieke, N., et al. (2020). The future of digital health with federated learning. npj Digital Medicine, 3, 119.
- Sheller, M. J., et al. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. Scientific Reports, 10, 12598.
- Kaissis, G. A., et al. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence, 2, 305–311.
- Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3–4), 211–407.
- Bonawitz, K., et al. (2017). Practical secure aggregation for privacy-preserving machine learning. ACM CCS.
- Li, T., et al. (2020). Federated optimization in heterogeneous networks. MLSys.
- Li, Q., et al. (2022). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 35(4), 3347–3366.
- Warnat-Herresthal, S., et al. (2021). Swarm learning for decentralized and confidential clinical machine learning. Nature, 594, 265–270.
- Geiping, J., et al. (2020). Inverting gradients—How easy is it to break privacy in federated learning? NeurIPS.
- Dayan, I., et al. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. Nature Medicine, 27, 1735–1743.