ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

# Gradient Without Borders: Homomorphic Encryption-Backed Collaborative Model Training Across Sovereign Healthcare Networks

**Mallesham Goli**
Independent Researcher
mallesham.goli.research01@gmail.com

## Abstract

The difficulties of working in sectors with strict privacy constraints, such as healthcare, result from the fact that data is separated and distributed across multiple sources that are unable to exchange it. The solution to this dilemma is to allow multiple data owners to collaborate on a joint task by using shared/global models while keeping the data decentralized. Three paradigms of federated machine learning are proposed: horizontal federated learning, vertical federated learning, and federated transfer learning. Federated learning processes and architectures with privacy-preserving properties that also address data provenance, access control, and interoperability.

Integrating heterogeneous distributed data sources while maintaining privacy is a major challenge for modern data and AI analysis. Sensitive data in particular, such as healthcare data, privacy is guaranteed by law. Therefore, many healthcare data sources are available based on these data sharing and data monetization are a difficult integration task. Multiple organizations are willing to collaborate by sharing insights from their private datasets, but without sharing the original datasets because healthcare data is highly sensitive. However, existing guidelines permit the use of healthcare data in Europe and America only for either research or health benefits, and not even for commercial use. Nevertheless, several companies are attempting to break this barrier by using fake-generated data for AI modeling purposes. As a result, three main criteria must be considered during the establishment of collaboration for multi-organization AI data modeling: privacy, security, and monitoring.

**Keywords :** Federated Learning,Privacy-Preserving Machine Learning,Secure Multi-Party Computation (SMPC),Healthcare Data Collaboration,Distributed AI Systems,Differential Privacy,Data Governance in Healthcare,Secure Model Aggregation,Cross-Institutional Data Analysis,Homomorphic Encryption,Decentralized Machine Learning,Clinical Data Sharing Frameworks,Privacy-Aware AI Architectures,Interoperable Healthcare Systems,Trustworthy AI in Healthcare.

## 1. Introduction

Healthcare data provide a large but underutilized resource for improving health. Security, privacy, and compliance challenges hinder sharing and federated learning can enable multi-organization collaboration without sharing data. Security concerns can also discourage participation even when there is no data leakage risk. Different federated learning architectures offer various degrees of protection against model inversion and membership inference attacks. Multi-party computation, differential privacy, and homomorphic encryption can be applied at different stages of the process to protect sensitive information, even when the model owner aims to recover it. Gaps in data governance, provenance handling, and interoperability standards represent additional risks to organisational and legal compliance. Despite these challenges, secure collaboration remains essential for developing AI models capable of responding to complex global health crises and enabling validation against real-world settings.

Federated AI technologies can offer effective solutions for these critical security, privacy, compliance, and usability issues associated with the secure collaboration of multiple organisations in the health sector. A comprehensive overview of the state of

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

the art was previously presented, with the focus on federated learning and secure multi-party computation paradigms for the analysis of federated healthcare data. The results of this broader analysis are reported in distinct sections addressing the well-established literature as well as recent advances.

Federated AI technologies provide promising solutions to the major security, privacy, compliance, and usability challenges that arise when multiple organizations collaborate in the healthcare sector. These technologies enable institutions to jointly analyze distributed healthcare data without requiring the direct sharing of sensitive patient information. In particular, federated learning allows machine learning models to be trained across decentralized datasets while keeping the data stored locally within each organization, thereby preserving privacy and regulatory compliance. Similarly, secure multi-party computation enables different parties to collaboratively compute results on encrypted data without revealing the underlying information to one another. A comprehensive overview of the current state of the art highlights how these approaches support secure and efficient collaboration across healthcare institutions. Previous analyses have examined both the well-established literature and the most recent advancements in federated AI technologies, providing insights into their capabilities, limitations, and practical applications. The findings from this broader review are organized into distinct sections that discuss foundational research as well as emerging developments in federated learning and secure computation techniques for federated healthcare data analysis.

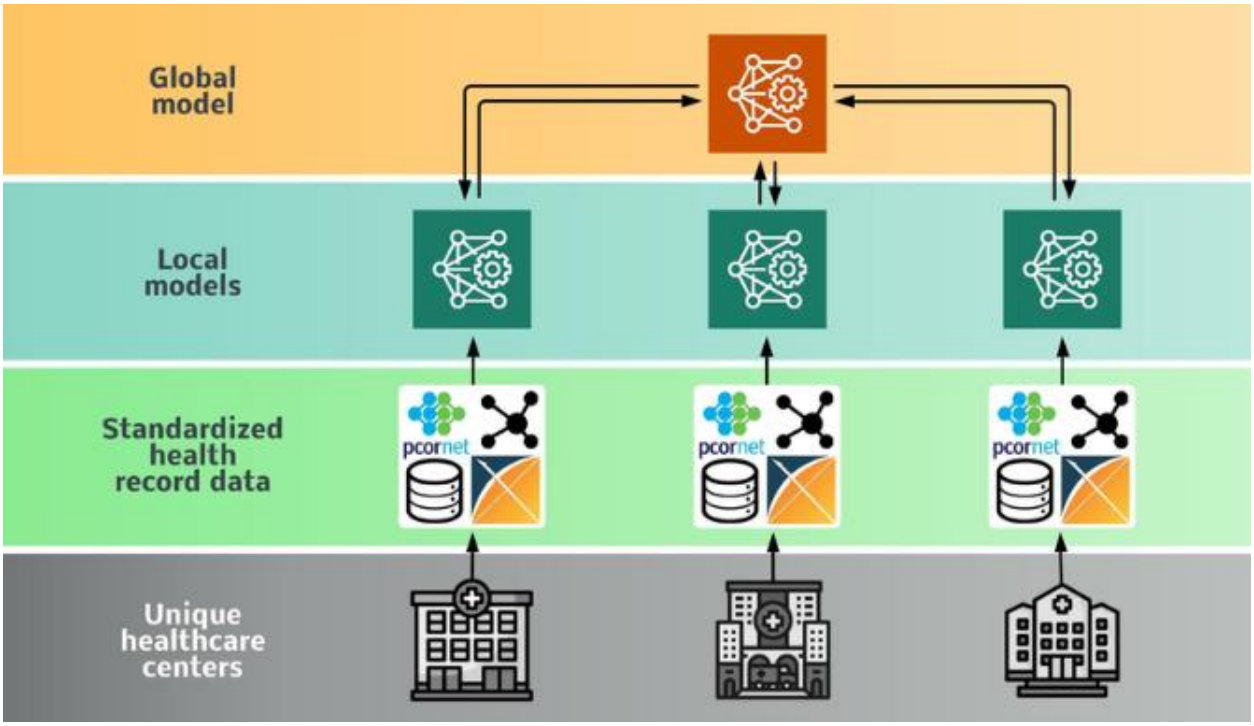



**Fig 1: Federated learning model architecture**

### 1.1. Background and Significance

In the digital world, cloud computing has become essential. It allows us to manage, analyze, and share large amounts of data

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

without needing to create our own IT infrastructure. Yet, most data are stored in clouds owned by companies that lack appropriate security and privacy measures, raising concerns in healthcare and finance. Private cloud implementations improve security but lack the collaborative capability of public clouds. Federated learning solves these issues by keeping sensitive data at their origin and running model training directly where the data are located. Organizations collaborating in federated learning are directly involved in the learning process. Paradigms are based on the amount of shared data. In horizontal federated learning, data with the same semantics are distributed among different nodes. In vertical federated learning, the same sample applies to different feature sets at different nodes. In a federated cross-silo setting, multiple organizations join forces to exploit performance at the generalization level by increasing the amount of training data.

Although federated learning paradigms provide models capable of supporting the detection of privacy issues in data associated with health and financial applications, the standard architectures for establishing federated learning failures require thorough analysis. Most implementations use a cloud model with third-party service providers acting as centralized orchestration. These cloud nodes receive all data for training but are not governed by involved authorities and thus are exposed. Using an architecture with a centralized third-party orchestrator exposes the final model to a non-governed organization. Analysis through the lens of privacy, security, and governance is crucial. When a cloud is used as a training mechanism, known and unknown attacks flow through the architecture, and adequate solutions must be defined.

**Equation 1: Federated learning (FL): objective → FedAvg update (step by step)**

- There are $K$ organizations (hospitals / labs / insurers).
- Organization $k$ has local dataset $D_k$ with size $n_k = | D_k |$.
- Total samples $N = \sum_{k=1}^{K} n_k$.
- Model parameters $w \in \mathbb{R}^d$.

Let per-example loss be $\ell(w; x_i, y_i)$. Local objective:

$$F_k(w) = \frac{1}{n_k} \sum_{i \in D_k} \ell(w; x_i, y_i).$$

A standard global objective consistent with "train across decentralized datasets" is:

$$F(w) = \sum_{k=1}^{K} p_k \; F_k(w), \text{where } p_k = \frac{n_k}{N}.$$

## 2. Background and Motivations

Federated Learning (FL) is a distributed machine learning (ML) scheme allowing algorithm training on non-IID sensitive data across organizations while maintaining data privacy. Compared to traditional distributed ML or parallel ML, FL offers

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

decentralized training without data sharing and enables edge-device-based federated transfer learning. Stand-alone FL is not sufficient for secure multi-organization data analysis; sensitive data storage and data-labeling tasks remain at risk when performed at a central organization. Orchestration-based FL emerges as a solution for these organizations.

Healthcare data analysis typically necessitates ML-based models trained on sensitive data from multiple organizations. Healthcare centers often collaborate on specific diseases, but sensitive data sharing is challenging due to unclear data ownership, lack of consent, and privacy regulations—especially in the EU, where healthcare data is considered sensitive. FL architectures for these analyses face three primary challenges: data ownership, preventive security for sensitive data storage, and orchestration governance. While requests to a central organization can be assessed in decision-making systems, FL data-assisted processes are more complex.

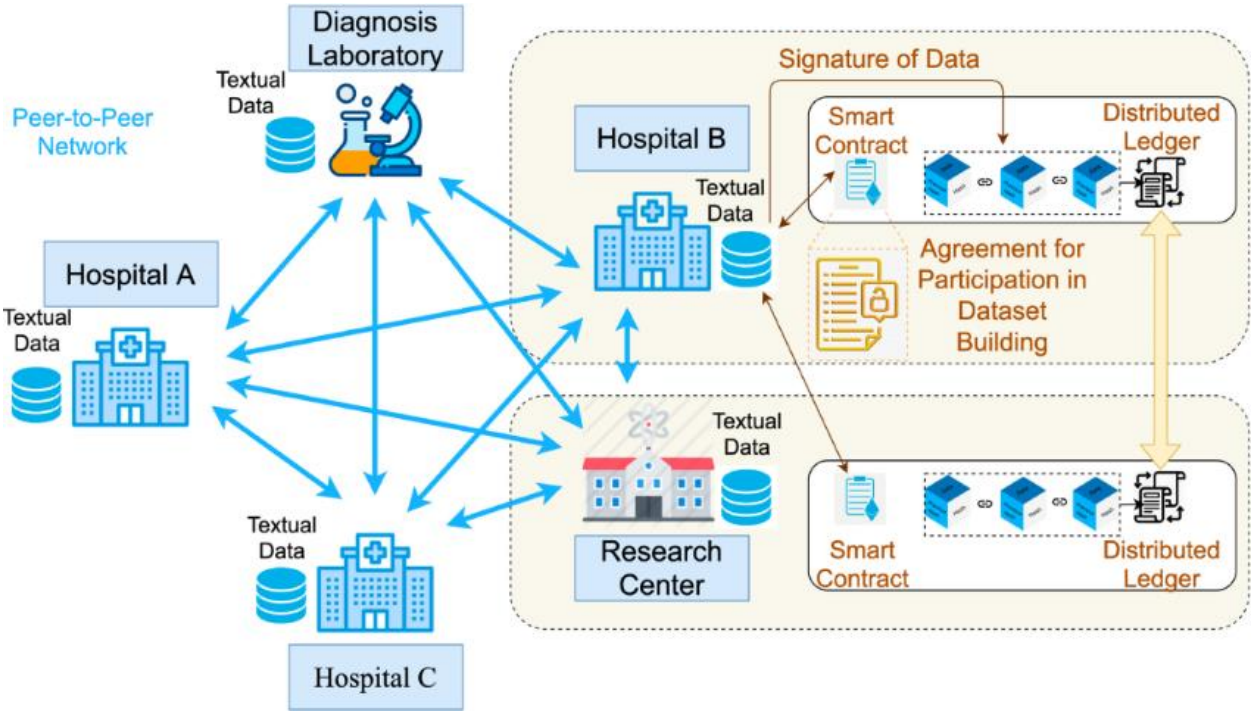



**Fig 2: Background and Motivations of Federated AI**

### 2.1. Research design

Recognizing that healthcare data contain a wealth of information needed for machine learning model training but that legal and ethical implications often make cross-organization data transfer difficult or impossible, federated learning enables organizations to collaboratively build ML models without sharing data. This large-scale systematic review identifies the current state of federated learning paradigms and AI architectures developed for secure collaborative healthcare data analysis. The discussion focuses on security, privacy, governance, and the compliance landscape, as well as exploring advanced privacy-preserving techniques such as homomorphic encryption, secure multi-party computation, and differential protection.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

The federated learning architecture reviewed allows multiple entities to collaboratively analyse their data while remaining compliant with regulations such as the European General Data Protection Regulation. Each entity holds a trusted execution environment, and deep learning models are generated by centralised model orchestration or decentralized aggregation. The federated learning paradigms used fully respect data privacy and ownership in federated environments such as the edge, cloud, and hybrid clouds, remaining robust against adversarial attacks.

**Equation 2: Local SGD: from gradient descent to local training**

Client $k$wants to reduce $F_k(w)$. Gradient descent step:

$$w \leftarrow w - \eta \nabla F_k(w).$$

But clients usually do **multiple local steps** $E$. Let $w_k^{t,0} = w^t$. For local step $e = 0, \dots, E-1$:

$$w_k^{t,e+1} = w_k^{t,e} - \eta\, g_k^{t,e},$$

where $g_k^{t,e}$is a stochastic gradient computed on a mini-batch $B_k^{t,e} \subset D_k$:

$$g_k^{t,e} = \frac{1}{\mid B_k^{t,e} \mid} \sum_{i \in B_k^{t,e}} \nabla_w \ell(w_k^{t,e}; x_i, y_i).$$

After $E$steps, client returns $w_k^{t,E}$to server.

Define the model delta (update) from client $k$:

$$\Delta_k^t = w_k^{t,E} - w^t.$$

The simplest federated averaging (FedAvg) aggregates client models (or deltas) using data-size weights:

$$w^{t+1} = \sum_{k \in S_t} \frac{n_k}{\sum_{j \in S_t} n_j} w_k^{t,E}.$$

Equivalently in delta form:

$$w^{t+1} = w^t + \sum_{k \in S_t} \frac{n_k}{\sum_{j \in S_t} n_j} \Delta_k^t.$$

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

## 3. Federated Learning Paradigms for Healthcare

Federated Learning Paradigms for Healthcare. The need for multi-party cooperation in healthcare data analytics can be met using a specific variant of distributed learning called federated learning. FL enables learning a shared model across multiple organizations while preserving the data privacy within the organizations. FL is broadly categorized into horizontal FL, vertical FL, and federated transfer learning.

Horizontal federated learning (HFL) is considered when multiple parties share the same feature space (the feature sets of the local datasets are the same) but with different samples (the sample space of the local datasets is different). A typical FL setup appears in supervised learning when each organization owns a local dataset that belongs to the same feature space. The goal of HFL is to train a global model that generalizes better than any local model while distributing the learning process without sharing the local datasets. The process of HFL is typically orchestrated by an external servicer that oversees the communication of the local machines with the server and integrates the local models to create an updated global model.

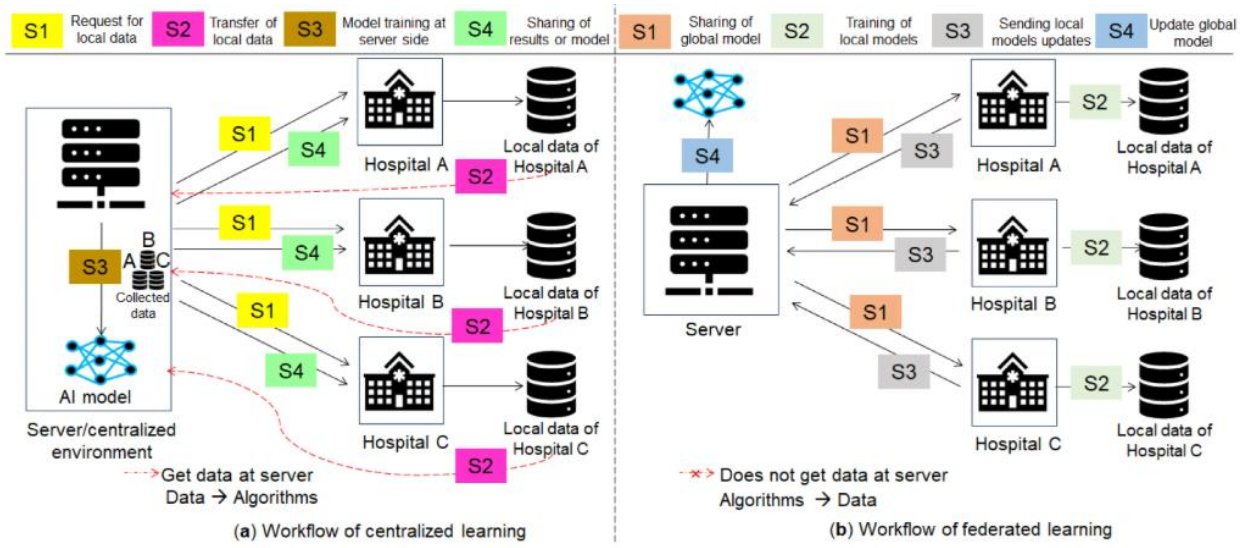



**Fig 3: Analysis of Federated Learning Paradigm in Medical Domain**

### 3.1. Horizontal Federated Learning

In horizontally partitioned data, each organization possesses samples representing the same data distribution but for different instances. A typical use case is when a federated learning system for healthcare applications is employed by a set of hospitals or clinical centers that collect samples for the same disease but represent their patients in different regions where external factors such as culture, diet, and weather may have an influence on the disease but not on the data distribution. An obvious system is the training of deep learning models for image disease classification based on data from different hospitals or clinical centers.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

Two main methods can be identified for horizontally partitioned data. The first method involves sharing the weights of the model with a central server, which performs only the aggregation task. The second method is for client-sending layers or model parameters. Although the two methods generated low regional accuracy in some applications, the overall accuracy using federated learning exceeded that of independent model training. Additionally, both horizontal partitioned data methods faster required less communication time than the non-federated learning model training method. Moreover, the systems experienced greater performance with a larger number of participating hospitals.

**Equation 3: Vertical FL (VFL): feature partitioning**

Same patient IDs (or partially overlapping), but features split:

$$x_i = [x_i^{(A)}, x_i^{(B)}], x_i^{(A)} \in \mathbb{R}^{d_A}, x_i^{(B)} \in \mathbb{R}^{d_B}, d_A + d_B = d.$$

Model parameters also split:

$$w = [w^{(A)}, w^{(B)}].$$

Now the **prediction** often needs both parties. For a linear/logistic model, define:

$$z_i = w^\top x_i = (w^{(A)})^\top x_i^{(A)} + (w^{(B)})^\top x_i^{(B)}.$$

## 4. Architectural Models for Secure Collaboration

Architectural models based on a centralized orchestrator appear to be the most effective for securely sharing healthcare data between multiple organizations. In these models, data is processed by second-party organizations using cryptographic techniques and sent to a third-party organization for storage and further use. The advantage of such models is that the third party can be responsible for computing-intensive tasks, such as training federation models.

Example architectures belonging to this category include SecureFL, which incorporates remodels of semi-honest third-party organizations and training malfeasant third-party organizations, Fed-implantable, which aims to cater to the needs of FDHC systems supported by a centralized third-party organization, and ProFL, which not only protects data privacy but also safeguards model confidentiality against the data provider. Analyzing a group of healthcare organizations under common control and regulation, a cross-domain federated softmax learning framework is proposed that considers both interaction and privacy-preserving requirements.

Another variant uses federated adaptive random selection to dynamically select training sites and hide their identity by secret sharing and local data perturbation. The model overcomes the inefficiency of the passive semi-honest third party and Byzantine-resilient heavy gradients in subsequent faceted learning. The proposed faceted deep learning framework ensures that patients can actively control, select, and share personal sensitive data with specific hospitals.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

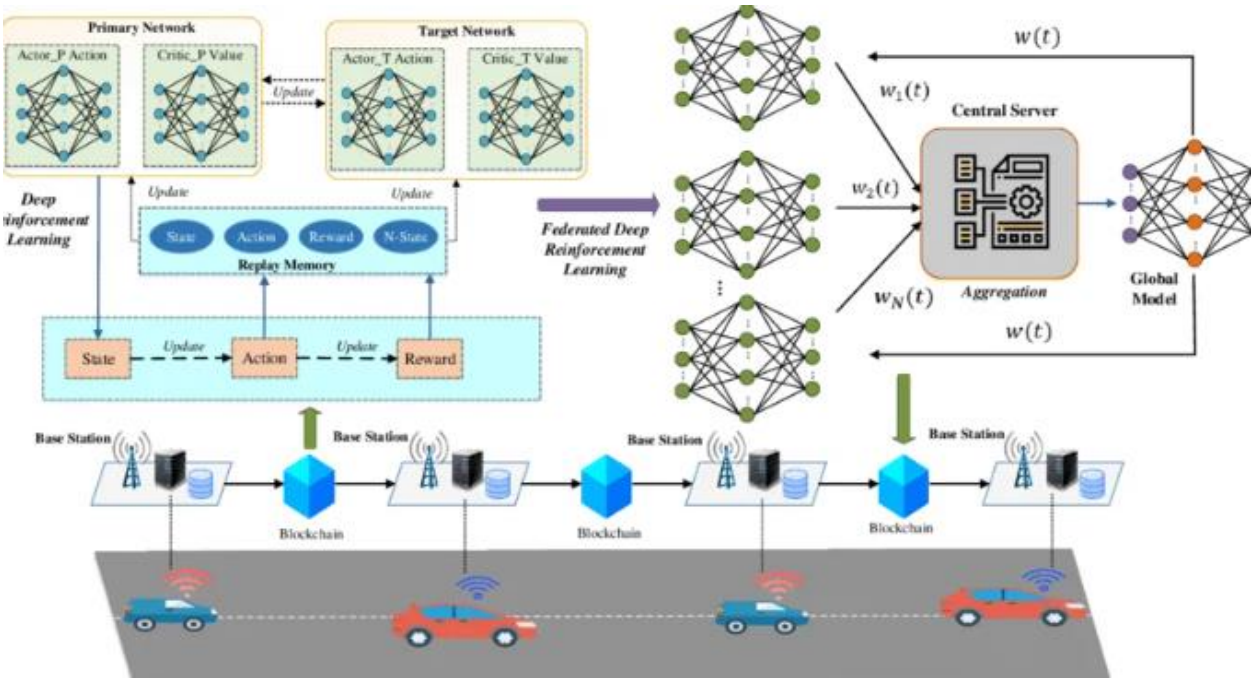



**Fig 4: The architecture of federated learning**

### 4.1. Centralized Orchestrator Architectures

An architecturally centralized federated learning model helps multi-organizational entities concerned with patient data analyses or private ML/AI solution offerings. The parties such as health providers, medical device manufacturers, health and pharmaceutical, insurance companies, supervised by a regulatory organization for the conducts have their own distributed dataset confined to their premises due to jurisdiction regulations and data-usage contracts. A central orchestrator allows independent multi-organization entities to participate in a federated paradigm that is centrally scrutinized and monitored, addressing some of the safety and reliability requirements needed for federated learning, such as patient safety, legal compliance of patient data usage, data governance, execution control, security, risk control, and thedata being the patients' assets.

An institution not gaining any benefits from the training, testing, or validation of services or products allowed to be use or offered with patient data can supervise the data-access contracts & processes of the data-consuming parties, the results gained from the analyses, and their distribution. The use of patient data by unauthorized parties or beyond the restrictions will be detected and reported by the supervisor. Once the usage of patient-data by these organizations is established, other institutions can readily join the federation, sharing & using ML solutions already trained by other agents in the association. The presence of one of the multi-organization sides (particularly health providers) as agent practicing model distinction is needed for patient safety during the supervised multi-organization federated analysis.

**Equation 4: Vertical FL example: logistic regression gradients with secure sum (full derivation)**

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

Assume binary logistic regression, label $y_i \in \{0,1\}$, probability:

$$\hat{y}_i = \sigma(z_i) = \frac{1}{1+e^{-z_i}}.$$

Cross-entropy loss:

$$\ell_i(w) = -y_i \log(\hat{y}_i) - (1-y_i)\log(1-\hat{y}_i).$$

A standard result (derive via chain rule) is:

$$\frac{\partial \ell_i}{\partial z_i} = \hat{y}_i - y_i.$$

Because $z_i = (w^{(A)})^\top x_i^{(A)} + (w^{(B)})^\top x_i^{(B)}$,

$$\nabla_{w^{(A)}} \ell_i = \frac{\partial \ell_i}{\partial z_i} \cdot \frac{\partial z_i}{\partial w^{(A)}} = (\hat{y}_i - y_i)\, x_i^{(A)}.$$

Similarly,

$$\nabla_{w^{(B)}} \ell_i = (\hat{y}_i - y_i)\, x_i^{(B)}.$$

## 5. Privacy, Security, and Compliance Considerations

Specific privacy or security techniques used to protect model training depend on the selected federated learning paradigm and the sensitivity of the data involved. Healthcare data is by default sensitive, requiring careful consideration of both data privacy and model integrity to ensure compliance with applicable regulations. At a conceptual level, the relevant approach is based on data minimization: sharing models rather than the underlying data violates neither security nor privacy.

Privacy provisioning techniques can be classified into three categories: model level, communication level, and data level [59]. Techniques acting at the model level can be roughly categorized according to whether privacy or security of the model itself is considered. Privacy-preserving data sharing, privacy-preserving data mining, and privacy-preserving machine learning belong to the former class. Homomorphic encryption (HE) and secure multi-party computation (SMPC) are typical techniques for the latter task. Techniques acting at the communication level aim to protect model updates; homomorphic encryption, differential privacy (DP), secret sharing, and model perturbation are all widely applied techniques. Finally, data-level techniques focus on the data used for training, and HE, SMPC, DP, feature selection, mulit-party computation, data encryption, and data perturbation are typical candidates.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

Homomorphic encryption, secure multi-party computation, and differential privacy are now treated in more detail to illustrate the specifics of the techniques involved. Further information can be found in survey articles on privacy-preserving techniques for federated learning.

**Equation 5: Secure sum via additive homomorphism (HE) (equation)**

Federated AI Architectures for …

If Enc is additively homomorphic:

$$\text{Enc}(a) \oplus \text{Enc}(b) = \text{Enc}(a + b).$$

Protocol sketch:

1. Party A computes $z_i^{(A)}$, sends $\text{Enc}(z_i^{(A)})$to B.
2. Party B computes $z_i^{(B)}$and combines:

$$\text{Enc}(z_i) = \text{Enc}(z_i^{(A)}) \oplus \text{Enc}(z_i^{(B)}) = \text{Enc}(z_i^{(A)} + z_i^{(B)}).$$

3. Authorized decryptor gets $z_i$, computes $\hat{y}_i$, returns only what's necessary (often masked/limited) for gradient steps.

The paper also lists SMPC.

Federated AI Architectures for …

Simple additive secret sharing over a large modulus $q$:

- A secret $s$is split into shares $s_1, \dots, s_m$such that

$$s \equiv \sum_{j=1}^{m} s_j \ (\text{mod} q).$$

**5.1. Data Privacy Techniques (HE, SMPC, DP)**

HE, SMPC, and differential privacy (DP) are the main data privacy techniques integrated into machine learning algorithms in real-world applications for secure data analysis in federated learning systems. HE empowers partner organizations to keep sensitive data private and encrypted. It allows computing on ciphertexts, generating an encrypted result that remains unintelligible until decryption. Thus, organizations can retain the confidentiality of valuable sensitive data in the unique or extended database during the federated learning process.

SMPC enables two parties to jointly evaluate a function while keeping their inputs private. Distributing the data across partner organizations and processing data without sharing it is a powerful approach to solving data-sharing and data-sensitivity issues.

ISSN: 3067-4166

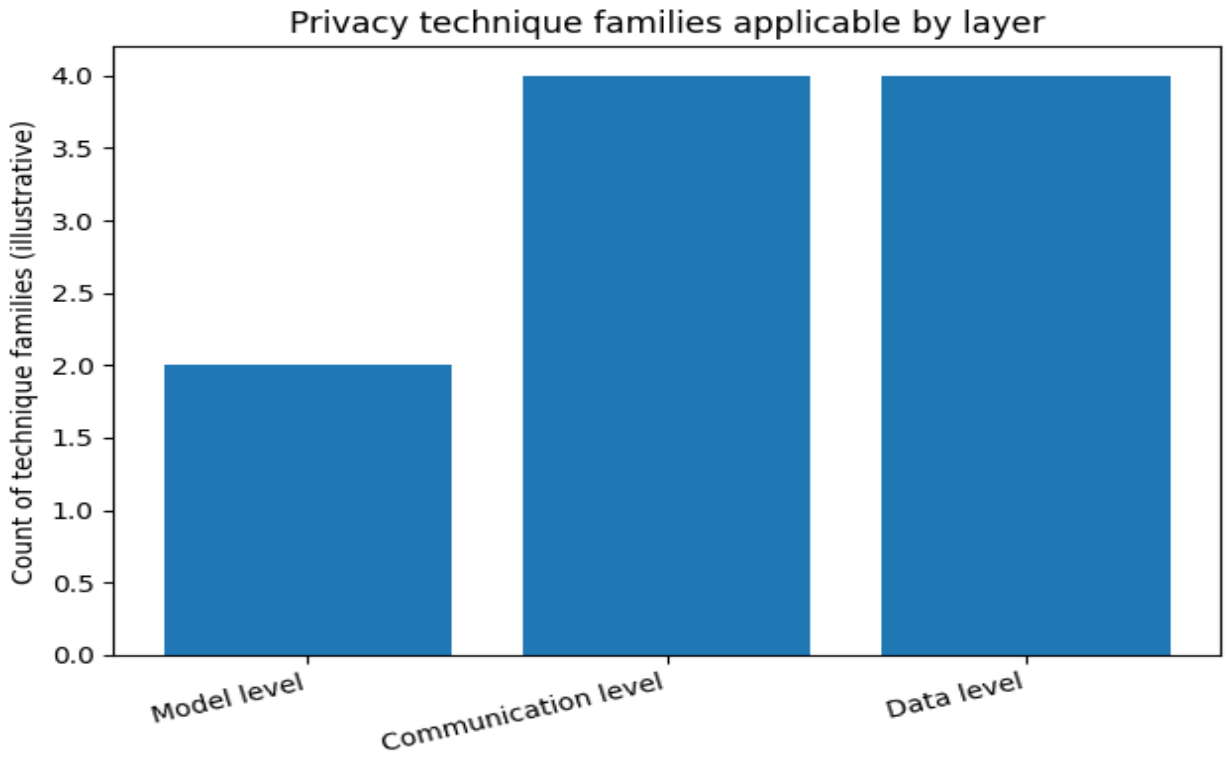
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

Differential privacy is achieved by adding noise to the data set in an appropriate way and removing it when the final result is obtained. When the outputs are published, neither the mechanism nor the statistician can infer anything about the input.

Data privacy techniques are critical in horizontal and vertical federated learning. However, a detailed discussion of the benefits, advantages, and limitations of such techniques in vertical or multi-party FL is beyond the scope of this work.




## 6. Data Governance, Interoperability, and Standards

Data governance aims to answer crucial questions regarding data management, such as possession and acceptable usage, including techniques for auditability, provenance, and access control. These aspects are especially relevant when the data owner is not the user, a typical situation in data-sharing or federated settings. Consequently, data governance is directly related to data privacy and its regulations: without provenance, auditability, and defined usage policies regarding access and trade, it is difficult, if not impossible, to comply with data privacy rules.

Data interoperability allows different systems, organizations, and people to exchange and mutually use data. It especially highlights the syntactic and semantic interoperability enabled by the use of common frameworks, protocols, and standards in the

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

definition of data description and access templates. Nonetheless, achieving semantic interoperability still requires additional effort, even with common definitions, due to differences in the data models used in data sources. Such differences may arise from the underlying technologies (for example, relational databases and blockchain), the specific domain (for example, healthcare and finance), or heterogeneous analysis perspectives (for example, prediction and what-if scenarios). While these concerns often do not apply in the same organization, they emerge when data are shared with third parties. Finally, data standards constitute a necessary condition—not sufficient, though—for data interoperability. Without common frameworks and protocols, achieving interoperability across different domains requires customizing connections for each pair of data sources.

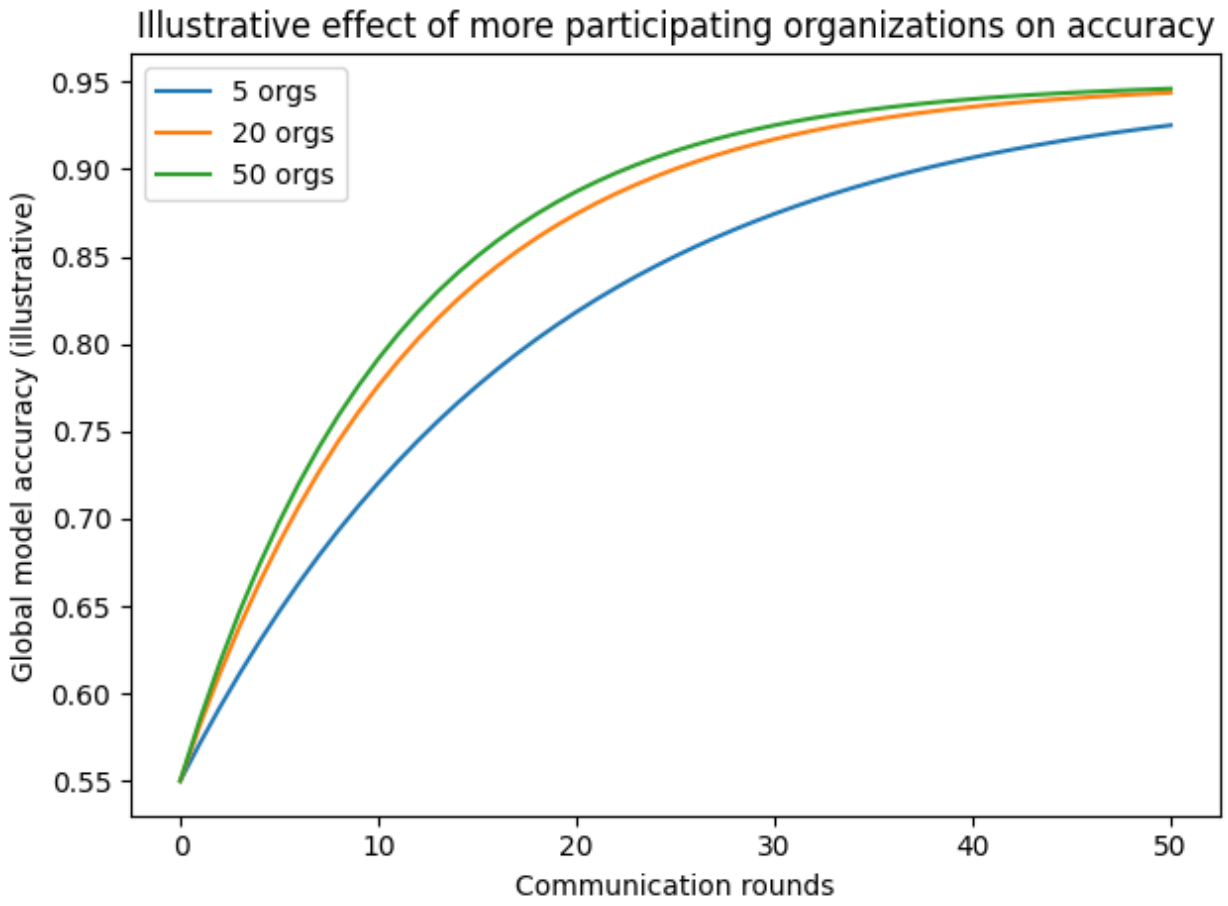
Data interoperability refers to the ability of different systems, organizations, and individuals to exchange data and use it effectively across various platforms. It relies heavily on **syntactic and semantic interoperability**, which are supported by shared frameworks, protocols, and data standards that define how data is structured and accessed. However, achieving full semantic interoperability remains challenging because data sources often rely on different **data models and technologies**, such as relational databases or blockchain systems. Variations may also arise from differences in application domains—like healthcare or finance—or from diverse analytical approaches, such as predictive modeling or scenario analysis. While these issues may be less prominent within a single organization, they become more significant when data is exchanged with external parties. Therefore, although data standards are essential for enabling interoperability, they alone are not sufficient. Without common frameworks and protocols, organizations must create customized integrations between each pair of data sources, making cross-domain data sharing more complex and resource-intensive.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18




**6.1. Data Provenance and Access Control**

Data provenance generated by the collaboration in federated learning demonstrates the quality and trustworthiness of the trained models beyond model performances. Provenance can be used to identify the participating parties and the data used to generate the models. By exposing sufficiently detailed provenance information, the participating organizations can share their knowledge–i.e., the knowledge about what is being modeled and the data provenance–with any external users of the models, for example, regulatory organizations. Policy-based provenance access control can restrict any sensitive provenance information from being exposed to untrusted users. In healthcare scenarios, provenance information can be sensitive because it provides knowledge about the data hosting organizations and their patient populations, which potential attackers can use for future data breaches, for example, identity theft.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

In the cyber security domain, the copyright of the used data and who provided the data for modeling can be extremely sensitive. Therefore, defining a security model for provenance access control can help reduce the risk of exposing sensitive provenance information to unauthorized users.

## 7. Conclusion

Great efforts have been made to explore FL in the sensitive domain of healthcare, which has huge potential for a wide variety of real use cases. However, almost all approaches consider a single organization setting, while in practice, many use cases require collaboration among multiple organizations to gain insights. Therefore, recent studies have provided federated FL paradigms for integrating FL with other distributed learning paradigms. Still, security, privacy, and governance aspects of multi-organizations collaboration remain largely unexplored. The proposed discussion has focused on these aspects while building a design strategy for the architectures of multi-organizations collaboration. The discussion is pivotal to initiate an in-depth analysis of multi-organization FL frameworks in the sensitive healthcare domain. Several critical aspects to achieve secure multi-organization collaboration using FL have been investigated, including multiple federated learning paradigms, architectural models of secure collaboration, privacy techniques on sensitive data, and data governance for enabling and sustaining collaboration.

The progress of FL is fast, but it is still in its infancy for a multi-organization setting. Many new FL paradigms have been introduced for enabling collaboration among multiple organizations. The concept of a central operator has also been extended to permit the sharing of local models while removing dependence on a central operator. Even with these substantial developments, security and privacy aspects of collaboration at the organizational level have rarely been studied, including clarifying the notion of control over the shared sensitive data and models. In particular, there has been little interaction with privacy techniques such as HE and SMPC, which are mainstream in secure data sharing at the organization level. No exploration has taken place into governance issues related to data provenance, control over sharing during actual FHE, and the relationships with traditional data governance topics in the multi-organization context.

| **Paradigm** | **What is shared?** | **Data partitioning** | **Typical healthcare example** |
|---|---|---|---|
| Horizontal FL (HFL) | Model updates / weights | Same features, different samples | Multi-hospital imaging classifier |
| Vertical FL (VFL) | Encrypted gradients / partial embeddings | Same samples (overlap), different features | Hospital + insurer joint risk model |
| Federated Transfer Learning (FTL) | Shared representations / transferred parameters | Different samples and features (limited overlap) | Rare disease model across countries |

**Table : Federated learning paradigms (summary table)**

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

**7.1. Future Trends**

Key enabling technologies for future healthcare service delivery, along with government policies and funding, are likely to accelerate market development. AI-powered solutions can assist users in identifying, anticipating, and resolving potential issues regarding health and telemedicine services. Better data sharing and collaboration among European healthcare institutions will also enhance the personalization of various preventive measures. For example, using federated learning makes it possible to train cluster models that better inform personalized care interactions with patients and improve the recommendation of personalized prevention actions.

Changes in the governance of current care delivery systems will also stimulate growth. In the long term, the COVID-19 pandemic is likely to improve risk management processes around healthcare systems worldwide. All nations are expected to invest in systems designed to ensure rapid, real-time data collection and analysis in order to facilitate timely decision-making and increase liquid capital. More accessible data structures will enable integrated operations across different systems and countries. New advanced analytics and predictive models, including AI-driven solutions, can also enhance user monitoring, allowing both patients and insurance companies to act preemptively and effectively prevent risk situations.

## 8. References

1. Guntupalli, R. (2025). Multi-Cloud vs. Hybrid Cloud Security: Key Challenges and Best Practices. Hybrid Cloud Security: Key Challenges and Best Practices (November 21, 2025).
2. Mell, P., & Grance, T. (2011). The NIST definition of cloud computing. NIST.
3. Sheelam, G. K. (2025). Agentic AI in 6G: Revolutionizing Intelligent Wireless Systems through Advanced Semiconductor Technologies. Advances in Consumer Research.
4. World Health Organization. (2021). Ethics and governance of artificial intelligence for health. WHO Press.
5. Kolla, S. H. (2024). RETRIEVAL-AUGMENTED GENERATION WITH SMALL LLMS FOR KNOWLEDGE-DRIVEN DECISION AUTOMATION IN ENTERPRISE SERVICE PLATFORMS. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(3), 476–486. https://doi.org/10.61841/turcomat.v15i3.15497
6. Moreau, L., et al. (2015). The W3C PROV family of specifications. Future Generation Computer Systems, 29(7), 161–165.
7. Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
8. Van Roy, P. (2009). Self-management in distributed systems. IEEE Computer, 42(12), 40–47.
9. Vardhan Kumar Bandi, V. D. (2024). Automated Feature Engineering Systems in Large-Scale Healthcare Data Environments. Journal of Neonatal Surgery, 13(1), 2127–2141. Retrieved from https://www.jneonatalsurg.com/index.php/jns/article/view/10004

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

10. Sutton, R. S. (2019). The bitter lesson. Incomplete Ideas Blog.
11. Floridi, L. (2013). The ethics of information. Oxford University Press.
12. Nigam, N., Sireesha, B., Ediga, P., Segireddy, A. R., & Bokde, S. (2025, December). Comparative Evaluation of Cloud Security Algorithms Using Multiple Classifiers with an Optimized Intrusion Detection System. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
13. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile Networks and Applications, 19, 171–209.
14. Pareyani, S., Goswami, S., Geetha, Y., Dimri, S. K., Niharika, D. S., & Amistapuram, K. (2025, December). Smart Resource Allocation in Wireless Sensor Networks Through AI Techniques. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
15. Dehghani, Z. (2022). Data mesh. O'Reilly Media.
16. Sriram, H. K., Gadi, A. L., & Challa, K. (2025). Leveraging AI, ML, and Gen AI in Automotive and Financial Services: Data-Driven Approaches to Insurance, Payments, Identity Protection, and Sustainable Innovation. Anil Lokesh and Challa, Kishore and singreddy, Sneha, Leveraging AI, ML, and Gen AI in Automotive and Financial Services: Data-Driven Approaches to Insurance, Payments, Identity Protection, and Sustainable Innovation (March 25, 2025).
17. Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3–4), 211–407.
18. Nagubandi, A. R. (2025). Cryptocurrency Market Spillovers: Risk Contagion Across Global Financial Systems.
19. European Parliament and Council of the European Union. (2016). General Data Protection Regulation (GDPR). Official Journal of the European Union.
20. Yandamuri, U. S. AI-Driven Decision Support Systems for Operational Optimization in Hospitality Technology.
21. Gentry, C. (2009). A fully homomorphic encryption scheme. Stanford University.
22. Guntupalli, R. (2025). Federated Deep Learning for Predictive Healthcare: A Privacy-Preserving AI Framework on Cloud-Native Infrastructure. Vascular and Endovascular Review, 8(16s), 200-210.
23. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
24. Dutta, P., Mondal, A., Vadisetty, R., Polamarasetti, A., Guntupalli, R., & Rongali, S. K. (2025). A novel deep learning rule-based spike neural network (SNN) classification approach for diagnosis of intracranial tumors. International Journal of Information Technology, 17(9), 5705-5712.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

25. He, J., Baxter, S., Xu, J., et al. (2019). The practical implementation of artificial intelligence technologies in medicine. Nature Medicine, 25, 30–36.
26. Enterprise-Scale Gen AI Orchestration Using Small LMs and LLM Agents for Intelligent ITSM and HRSD Automation in Enterprise Ecosystems. (2025). MSW Management Journal, 35(2), 1889-1897.
27. Holzinger, A. (2016). Interactive machine learning for health informatics. Springer.
28. FinOps Strategies for AI-Enabled Real-Time Compliance Platforms in Cloud Native Environments. (2025). MSW Management Journal, 35(2), 2080-2088.
29. IBM. (2023). Data fabric architecture overview. IBM Redbooks.
30. Velangani Divya Vardhan Kumar Bandi. (2024). Intelligent Data Platforms For Personalized Retail Analytics At Scale. Metallurgical and Materials Engineering, 30(4), 1011–1027. Retrieved from https://metall-mater-eng.com/index.php/home/article/view/1011-1027
31. Jennings, N. R., & Wooldridge, M. (1998). Applications of intelligent agents. Springer.
32. Sasi Kumar Kolla. (2023). Big Data–Driven Machine Learning Frameworks for Clinical Risk Prediction. International Journal of Medical Toxicology and Legal Medicine, 26(3 and 4), 44–59. Retrieved from https://ijmtlm.org/index.php/journal/article/view/1456
33. Kelly, C. J., Karthikesalingam, A., Suleyman, M., et al. (2019). Key challenges for delivering clinical impact with AI. BMC Medicine, 17, 195.
34. Kumar, K. M., Parasar, A., Walia, A., Inala, R., & Thulasimani, T. (2025, August). Enhancing Risk Management Strategies in Financial Institutions Using CNN and Support Vector Regression. In 2025 5th Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.
35. Koller, D., & Friedman, N. (2009). Probabilistic graphical models. MIT Press.
36. Rao, A. N., Garapati, R. S., Suganya, R. T., Kaliappan, A., & Kamaleshwar, T. (2025, August). Smart Solar Harvesting and Power Management in IoT Nodes Through Deep Learning Models. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.
37. Liu, F., et al. (2025). Foundational architecture for AI agents in healthcare. Cell Reports Medicine, 6(10), 102374.
38. Paleti, S., Baliyan, M., Aitha, A. R., Reddy, B. A., Bhadauria, G. S., & Sing, S. A. (2025, August). Graph—LSTM Hybrid Model for Improving Fraud Detection Accuracy in E-Commerce Financial Services. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.
39. Moreau, L., & Groth, P. (2013). Provenance: An introduction to PROV. Morgan & Claypool.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

40. Nagabhyru, K. C., Rani, M., Reddy, D. S., & Krishnaraj, V. (2025, August). Machine Learning-Driven Fault Detection in Electric Vehicles via Hybrid Reinforcement Learning Model. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.
41. Obermeyer, Z., & Emanuel, E. (2016). Predicting the future—Big data and clinical medicine. NEJM, 375, 1216–1219.
42. Vijaya Rama Raju Gottimukkala. (2025). Agentic AI for Next-Generation Cross-Border Payments: Contextual Learning in Transaction Routing. Journal of Informatics Education and Research, 5(4). Retrieved from https://jier.org/index.php/journal/article/view/3794
43. Pearl, J. (2009). Causality (2nd ed.). Cambridge University Press.
44. Srikanth, T., Segireddy, A. R., & Elavarasi, S. A. (2025, October). STaSFormer-SGAD: Semantic Triplet-Aware Spatial Flow-Guided Spatio-Temporal Graph for Anomaly Detection in Surveillance Videos. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1-7). IEEE.
45. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. NEJM, 380, 1347–1358.
46. Amistapuram, K. (2025). Agentic AI for Next-Generation Insurance Platforms: Autonomous Decision-Making in Claims and Policy Servicing. Journal of Marketing & Social Research, 2, 88-103.
47. Rieke, N., Hancox, J., Li, W., et al. (2020). Federated learning for digital health. NPJ Digital Medicine, 3, 119.
48. Challa, K., Sriram, H. K., & Gadi, A. L. (2025). Leveraging AI, ML, and Gen AI in Automotive and Financial Services: Data-Driven Approaches to Insurance, Payments, Identity Protection, and Sustainable Innovation.
49. Amershi, S., Begel, A., Bird, C., et al. (2019). Software engineering for machine learning: A case study. Proceedings of the International Conference on Software Engineering, 291–300.
50. Pamisetty, A., Paleti, S., Adusupalli, B., Singireddy, J., Inala, R., & Nagabhyru, K. C. (2025, September). Explainable AI Systems for Credit Scoring and Loan Risk Assessment in Digital Banking Platforms. In 2025 IEEE 13th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (pp. 1478-1483). IEEE.
51. Armbrust, M., Zaharia, M., Xin, R. S., et al. (2015). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59(11), 56–65.
52. Garapati, R. S. (2025). An Intelligent IoT Security System: Cloud-Native Architecture with Real-Time AI Threat Detection and Web Visualization. Journal homepage: https://jmsronline. com, 2(06).

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

53. Batini, C., & Scannapieco, M. (2016). Data and information quality: Dimensions, principles and techniques. Springer.
54. Babaiah, C., Dobriyal, N., Shamila, M., Aitha, A. R., Patel, S. P., & Upodhyay, D. (2025, December). Intelligent Fault Detection and Recovery in Wireless Sensor Networks Using AI. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
55. Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices. NPJ Digital Medicine, 3, 118.
56. Nagabhyru, K. C. (2025). Beyond Automation: The 2025 Role of Agentic AI in Autonomous Data Engineering and Adaptive Enterprise Systems.
57. Bertsekas, D. P. (2012). Dynamic programming and optimal control (Vol. 1). Athena Scientific.
58. Vajpayee, A., Khan, S., Gottimukkala, V. R. R., Sharma, D., & Seshasai, S. J. (2025). Digital Financial Literacy 4.0: Consumer Readiness for AI-Driven Fintech and Blockchain Ecosystems. International Insurance Law Review, 33(S5), 963-973.
59. Brundage, M., Avin, S., Clark, J., et al. (2018). The malicious use of artificial intelligence. arXiv.
60. Inala, R. (2025). A Unified Framework for Agentic AI and Data Products: Enhancing Cloud, Big Data, and Machine Learning in Supply Chain, Insurance, Retail, and Manufacturing. EKSPLORIUM-BULETIN PUSAT TEKNOLOGI BAHAN GALIAN NUKLIR, 46(1), 1614-1628.
61. Ferber, J. (1999). Multi-agent systems: An introduction. Addison-Wesley.
62. Garapati, R. S., & Daram, D. S. B. (2025). AI-Enabled Predictive Maintenance Framework For Connected Vehicles Using Cloud-Based Web Interfaces. Available at SSRN 5524261.
63. Kephart, J. O., & Chess, D. M. (2003). The vision of autonomic computing. Computer, 36(1), 41–50.
64. Aitha, A. R., & Jyothi Babu, D. A. (2025). Agentic AI-Powered Claims Intelligence: A Deep Learning Framework for Automating Workers Compensation Claim Processing Using Generative AI. Available at SSRN 5505223.
65. Huhns, M. N., & Singh, M. P. (1998). Readings in agents. Morgan Kaufmann.
66. Nagabhyru, K. C., & Babu, A. J. Human In The Loop Generative AI: Redefining Collaborative Data Engineering For High Stakes Industries.
67. Erl, T. (2016). Microservices design patterns. Prentice Hall.
68. Gottimukkala, V. R. R. (2025). Generative AI for Exceptions and Investigations: Streamlining Resolution Across Global Payment Systems. Journal of International Commercial Law and Technology, 6(1), 969-972.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

69. Fowler, M. (2018). Refactoring (2nd ed.). Addison-Wesley.
70. Segireddy, A. R. (2025). GENERATIVE AI FOR SECURE RELEASE ENGINEERING IN GLOBAL PAYMENT NETWORK. Lex Localis: Journal of Local Self-Government, 23.
71. Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). Design patterns. Addison-Wesley.
72. Amistapuram, K. (2025). GENERATIVE AI FOR CLAIMS EXCEPTIONS AND INVESTIGATIONS: ENHANCING RESOLUTION EFFICIENCY IN COMPLEX INSURANCE PROCESSES. Available at SSRN 5785482.
73. Zaharia, M., et al. (2010). Spark: Cluster computing with working sets. HotCloud.
74. Singireddy, S. (2025, May). AI-Driven Comprehensive Insurance and AAA Membership Benefits Overview. In 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE) (pp. 1-13). IEEE.
75. Lakshman, A., & Malik, P. (2010). Cassandra. ACM SIGOPS Operating Systems Review, 44(2), 35–40.
76. Nagubandi, A. R. (2025). PIONEERING SELF-ADAPTIVE AI ORCHESTRATION ENGINES FOR REAL-TIME END-TO-END MULTI-COUNTERPARTY DERIVATIVES, COLLATERAL, AND ACCOUNTING AUTOMATION: INTELLIGENCE-DRIVEN WORKFLOW COORDINATION AT ENTERPRISE SCALE. Lex Localis, 23(S6), 8598-8610.
77. Stonebraker, M., & Çetintemel, U. (2005). One size fits all? ICDE Proceedings, 2–11.
78. Yandamuri, U. S. (2022). Big Data Pipelines for Cross-Domain Decision Support: A Cloud-Centric Approach. International Journal of Scientific Research and Modern Technology, 227.
79. Moreira, M. W. L., et al. (2018). IoT-based smart healthcare systems. Sensors, 18(4), 1155.
80. Challa, K. (2025). Innovations in Digital Finance and Intelligent Technologies: A Deep Dive into AI, Machine Learning, Cloud Computing, and Big Data in Transforming Global Payments and Financial Services. Deep Science Publishing.
81. Lebcir, I., Mageswari, S. U., Bhosale, Y. H., Nagubandi, A. R., & Mahabooba, M. M. Agile Strategic Management in the Age of Disruption: Leveraging AI and Data Analytics for Competitive Advantage.
82. Kumar, S. S., Singireddy, S., Nanan, B. P., Recharla, M., Gadi, A. L., & Paleti, S. (2025). Optimizing edge computing for big data processing in smart cities. Metallurgical and Materials Engineering, 31(3), 31-39.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

83. Yandamuri, U. S. (2023). An Intelligent Analytics Framework Combining Big Data and Machine Learning for Business Forecasting. International Journal Of Finance, 36(6), 682-706.
84. Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3), 160.
85. GUNTUPALLI, R. (2025). EXPLAINABLE AI IN CLINICAL DECISION SUPPORT: INTERPRETABLE NEURAL MODELS FOR TRUSTWORTHY HEALTHCARE AUTOMATIONEXPLAINABLE AI IN CLINICAL DECISION SUPPORT: INTERPRETABLE NEURAL MODELS FOR TRUSTWORTHY HEALTHCARE AUTOMATION. TPM–Testing, Psychometrics, Methodology in Applied Psychology, 32(S9 (2025): Posted 15 December), 462-471.
86. Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical decision support in the era of AI. JAMA, 320(21), 2199–2200.
87. Annapareddy, V. N., Challa, K., Komaragiri, V. B., Sriram, H. K., Kalisetty, S., & Goma, T. (2025, April). Explaining AI Techniques such as SHAP LIME and RISE in Limited Sample Size Neuroimaging Studies. In 2025 4th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1-6). IEEE.
88. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning (2nd ed.). MIT Press.
89. Siva Hemanth Kolla. (2023). Deep Learning–Driven Retrieval-Augmented Generation for Enterprise ITSM Automation: A Governance-Aligned Large Language Model Architecture . Journal of Computational Analysis and Applications (JoCAAA), 31(4), 2489–2502. Retrieved from https://www.eudoxuspress.com/index.php/pub/article/view/4774
90. Tsamados, A., Aggarwal, N., Cowls, J., et al. (2022). The ethics of algorithms. AI & Society, 37, 215–230.
91. Davuluri, P. S. L. N. . (2024). AI-Driven Data Governance Frameworks for Automated Regulatory Reporting and Audit Readiness. Metallurgical and Materials Engineering, 30(4), 996–1010. Retrieved from https://metall-mater-eng.com/index.php/home/article/view/1936
92. Wooldridge, M. (2009). An introduction to multiagent systems (2nd ed.). Wiley.
93. Bandi, V. D. V. K. (2023). Production-Grade Machine Learning Pipelines For Healthcare Predictive Analytics. South Eastern European Journal of Public Health, 189–205. Retrieved from https://www.seejph.com/index.php/seejph/article/view/7057
94. Zhang, A., Xing, L., Zou, J., & Wu, J. C. (2022). Shifting ML for healthcare to deployment. Nature Biomedical Engineering, 6, 1330–1345.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 02 REVISED: JANUARY 19 ACCEPTED: FEBUARY 27 PUBLISHED: MARCH 18

95. Kolla, S. K. (2021). Architectural Frameworks for Large-Scale Electronic Health Record Data Platforms. Current Research in Public Health, 1(1), 1–19. Retrieved from https://www.scipublications.com/journal/index.php/crph/article/view/1372
96. McMahan, B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 1273–1282.
97. Annapareddy, V. N., Singireddy, J., Preethish Nanan, B., & Burugulla, J. K. R. (2025). Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling. Jai Kiran Reddy, Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling (April 14, 2025).
98. Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports, 10*(1), 12598.
99. Nandan, B. P. (2022). AI-Powered Fault Detection In Semiconductor Fabrication: A Data-Centric Perspective.
100. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H., Albarqouni, S., Bakas, S., Galtier, M., Landman, B., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R., Trask, A., Xu, D., Baust, M., & Cardoso, M. (2020). The future of digital health with federated learning. *npj Digital Medicine, 3*(1), 119.
101. Sheelam, G. K., & Nandan, B. P. (2022). Integrating AI And Data Engineering For Intelligent Semiconductor Chip Design And Optimization. Migration Letters, 19, 2178-2207.
102. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine, 37*(3), 50–60.