Federated Learning Architectures for Privacy-Preserving Multi-farm Data Sharing: A Comprehensive Framework for Secure Collaborative Intelligence in Precision and Digital Agriculture Systems

Dr. Martijn Willem Jansen [1], Saskia Maria Brouwer [2], Matthew Scott Walker [3*]
[1] Professor, Department of Biosystems Engineering and Agricultural AI, Wageningen University & Research, Netherlands
[2] Professor, Digital Agriculture and Smart Farm Networks Unit, Wageningen University & Research, Netherlands
[3] Centre for Digital Agriculture and Secure Data Systems, University of Sydney, Australia

* Corresponding Author: **Matthew Scott Walker**

Article Info

P-ISSN: 3051-3421
E-ISSN: 3051-343X
Volume: 06
Issue: 01
January - June 2025
Received: 01-01-2025
Accepted: 02-02-2025
Published: 03-03-2025
Page No: 34-42

Abstract
The rapid expansion of precision agriculture technologies has generated vast quantities of heterogeneous farm-level data, yet traditional centralized data-sharing models face significant barriers including privacy concerns, data sovereignty issues, competitive sensitivities, and regulatory compliance challenges that limit collaborative intelligence development across multiple agricultural stakeholders. Federated learning emerges as a transformative paradigm that enables secure multi-farm collaboration by training shared machine learning models on distributed datasets without requiring raw data exchange, thereby preserving individual farm privacy while enabling collective intelligence. This review examines federated learning frameworks specifically designed for secure agricultural data sharing, analyzing core architectures including horizontal, vertical, and federated transfer learning approaches adapted for farm environments. We explore essential privacy-preserving mechanisms such as secure aggregation protocols, differential privacy techniques, homomorphic encryption, and blockchain-based trust management that protect sensitive farm information during collaborative model training. Key agricultural applications are discussed, including federated crop yield prediction, distributed disease surveillance systems, cross-farm pest forecasting, and collaborative decision-support platforms that leverage multi-stakeholder data while maintaining data ownership. Despite promising advances, significant challenges remain regarding communication efficiency in rural networks, heterogeneity across diverse farming systems, model convergence with non-identically distributed data, and establishment of governance frameworks for federated agricultural intelligence. Successful deployment of federated learning in multi-farm contexts requires interdisciplinary integration of distributed machine learning, agricultural domain knowledge, cybersecurity protocols, and stakeholder engagement strategies to realize scalable privacy-preserving collaborative intelligence for sustainable digital agriculture transformation.

DOI: https://doi.org/10.54660/JADR.2025.6.1.34-42

Keywords: Federated Learning, Privacy-Preserving Machine Learning, Multi-Farm Data Sharing, Secure Aggregation, Digital Agriculture, Distributed Intelligence

Introduction

Modern agricultural systems increasingly rely on data-driven decision-making enabled by precision agriculture technologies, Internet of Things sensors, satellite imagery, unmanned aerial vehicles, and farm management information systems that continuously generate detailed information about soil conditions, crop health, weather patterns, management practices, and productivity outcomes [1, 2]. This data revolution presents unprecedented opportunities for developing sophisticated machine learning models that can optimize resource utilization, predict yields, detect diseases early, and support climate adaptation

strategies across diverse agricultural contexts [3, 4]. However, individual farms typically possess limited datasets that may not capture the full spectrum of environmental variability, crop genetics, pest pressures, and management scenarios necessary to train robust predictive models with adequate generalization capabilities [5].

Collaborative data sharing across multiple farms could substantially enhance model performance by aggregating diverse experiences and expanding training datasets to encompass broader ecological and operational conditions [6, 7]. Unfortunately, traditional centralized data-sharing approaches face formidable obstacles in agricultural contexts. Farmers and agricultural enterprises exhibit legitimate concerns about protecting proprietary information related to management practices, productivity levels, input costs, and strategic planning that could affect competitive positioning [8]. Additionally, regulatory frameworks including data protection legislation impose strict requirements on agricultural data handling, particularly when information could identify individual operations or reveal commercially sensitive insights [9]. Data sovereignty considerations, where farmers retain ownership and control over information generated on their properties, further complicate centralized aggregation models [10].

These challenges have motivated exploration of privacy-preserving collaborative learning paradigms that enable knowledge extraction from distributed datasets without requiring raw data centralization [11, 12]. Federated learning represents a particularly promising framework that allows multiple farms to jointly train shared machine learning models while keeping sensitive data localized on individual farm systems [13, 14]. By transmitting only model updates rather than raw observations, federated learning architectures can preserve privacy, maintain data sovereignty, and reduce communication overhead compared to conventional centralized approaches [15].

This article provides a comprehensive review of federated learning frameworks designed specifically for secure multi-farm data sharing in digital agriculture contexts. We examine fundamental architectures, privacy-preserving mechanisms, agricultural applications, and implementation challenges that characterize this emerging field. The objective is to synthesize current knowledge, identify research gaps, and provide guidance for developing practical federated learning systems that can accelerate collaborative agricultural intelligence while respecting privacy, security, and governance requirements.

Scope and Organization

Following this introduction, the article examines federated learning fundamentals adapted for agricultural contexts, analyzes security and privacy-preserving technologies essential for multi-farm collaboration, reviews agricultural applications that benefit from federated approaches, and discusses challenges and future perspectives for deploying these systems at scale. Throughout, emphasis is placed on practical considerations for implementing federated learning in real-world farm environments rather than generic machine learning or cybersecurity discussions.

Federated Learning Fundamentals

Core Principles and Architectural Paradigms

Federated learning fundamentally restructures the machine learning workflow by inverting traditional centralized training paradigms [16, 17]. Rather than aggregating data at a central location for model training, federated learning distributes the training process itself to data sources, enabling model development on decentralized datasets while maintaining data locality [18]. In agricultural contexts, this means machine learning models are trained directly on farm management systems, edge computing devices, or local servers where agricultural data originates, eliminating the need to transfer sensitive farm information to external repositories [19].

Three primary federated learning architectures have relevance for multi-farm agricultural applications. Horizontal federated learning applies when participating farms collect similar types of features but represent different sample populations [20]. For example, multiple crop farms monitoring soil moisture, temperature, precipitation, and management practices could collaboratively train yield prediction models using horizontal federation, where each farm contributes observations with identical feature structures but different data instances. Vertical federated learning addresses scenarios where different stakeholders possess complementary features about overlapping entities [21]. Agricultural supply chains might employ vertical federation where farms provide production data, weather services contribute meteorological features, and agricultural retailers supply input information to collaboratively predict outcomes for shared geographic regions. Federated transfer learning enables knowledge transfer between related but distinct agricultural domains, allowing models trained on data-rich farming systems to be adapted for data-scarce contexts with different crop types or environmental conditions [22].

Training Workflows and Communication Protocols

The federated learning process typically follows an iterative workflow coordinated by a central aggregation server that orchestrates model training without accessing raw farm data [23]. Initially, the server distributes a global model to participating farms, which then train this model locally using their respective datasets through standard gradient descent optimization procedures. After completing local training epochs, farms compute model updates representing the difference between initial and locally-trained model parameters. These updates, rather than raw data, are transmitted to the aggregation server, which combines contributions from multiple farms using aggregation algorithms to produce an improved global model [24]. This updated global model is redistributed to farms, and the process iterates until convergence criteria are satisfied.

Communication efficiency constitutes a critical consideration for agricultural federated learning, particularly given limited connectivity in rural environments [25]. Several strategies reduce communication overhead, including gradient compression techniques that transmit only significant parameter updates, federated averaging algorithms that decrease communication frequency by performing multiple local training epochs between aggregation rounds, and asynchronous protocols that accommodate farms with varying computational capabilities or intermittent connectivity [26, 27].

Comparison with Alternative Learning Paradigms

Federated learning occupies a distinct position within the spectrum of distributed machine learning approaches. Unlike fully centralized learning where all data resides in a single

repository, federated learning maintains data distribution, providing inherent privacy advantages and respecting data sovereignty principles essential for agricultural stakeholders [28]. Compared to completely decentralized peer-to-peer learning without coordination, federated learning employs orchestrated aggregation that typically achieves superior model convergence and consistency. Split learning represents an alternative approach where model layers are partitioned between clients and servers, but requires more communication rounds and may not provide equivalent privacy guarantees for agricultural applications [29].

Secure and Privacy-Preserving Multi-farm Learning

Secure Aggregation and Encryption Mechanisms

While federated learning prevents raw data sharing, model updates transmitted during aggregation can potentially leak sensitive information about individual farm datasets through model inversion attacks or gradient analysis [30]. Secure aggregation protocols address these vulnerabilities by ensuring the central server learns only the aggregate of all farm contributions without accessing individual farm updates [31]. Cryptographic techniques including secret sharing distribute each farm's model update across multiple servers such that no single entity can reconstruct individual contributions, while the aggregate remains computable. Homomorphic encryption allows mathematical operations on encrypted model updates, enabling aggregation without decryption, though computational overhead remains a practical limitation for resource-constrained farm systems [32, 33].

Secure multi-party computation protocols extend these protections by enabling collaborative computations where participating farms jointly compute functions over their collective inputs while learning only the final result [34]. For agricultural applications, this enables farms to collaboratively calculate benchmarks, identify optimal practices, or detect regional disease patterns without revealing individual farm performance or disease incidence to other participants.

Differential Privacy and Noise Injection

Differential privacy provides mathematical guarantees that individual farm contributions cannot be distinguished within aggregated models, protecting against inference attacks that might extract farm-specific information from trained models [35, 36]. By injecting carefully calibrated noise into model updates or gradients before aggregation, differential privacy mechanisms ensure that removing or modifying any single farm's data produces statistically indistinguishable model outcomes. The privacy budget parameter controls the trade-off between privacy protection strength and model accuracy, requiring careful calibration for agricultural applications where both privacy and prediction performance are essential [37].

Local differential privacy applies noise injection at individual farms before transmission, providing protections even if the aggregation server is compromised, though typically requiring larger noise magnitudes that may degrade model quality. Central differential privacy adds noise during aggregation, achieving better privacy-utility trade-offs when the central server is trustworthy. Adaptive noise mechanisms that account for heterogeneous data distributions across farms represent an active research area particularly relevant for agricultural contexts characterized by substantial environmental and operational variability.

Trust Management and Blockchain Integration

Multi-farm federated learning requires trust frameworks that verify participant identities, ensure model update authenticity, prevent malicious contributions that could poison collaborative models, and maintain transparent audit trails. Blockchain technologies offer decentralized trust infrastructure particularly suited for agricultural federated learning, where mutual trust among competing farms may be limited. Smart contracts can automate federated learning workflows, enforcing participation rules, validating model updates according to predefined quality criteria, and distributing incentives or access rights based on contribution levels.

Blockchain-based federated learning architectures store aggregation results, participation records, and model versioning information on distributed ledgers, creating immutable audit trails that support accountability and governance. Consensus mechanisms ensure that model updates meet quality standards before incorporation into global models, protecting against adversarial farms attempting to bias or degrade collaborative intelligence. Privacy-preserving blockchain protocols combine cryptographic techniques with distributed ledger benefits, maintaining confidentiality while providing transparency regarding participation and contribution patterns.

Data Ownership and Governance Frameworks

Effective multi-farm federated learning requires clear governance structures defining data ownership, usage rights, intellectual property in collaboratively-trained models, and benefit-sharing mechanisms. Farms typically retain ownership of raw data residing on their systems, while governance frameworks must address ownership and access rights for derived models trained using collective contributions. Licensing agreements, data cooperatives, and consortium arrangements provide organizational structures through which farms can establish collaborative learning agreements while protecting individual interests.

Federated learning governance should specify participant eligibility criteria, contribution requirements, model access policies, quality assurance procedures, dispute resolution mechanisms, and exit provisions. Particular attention is required for ensuring equitable participation opportunities for farms with varying technological capacities, preventing situations where technically sophisticated operations dominate collaborative intelligence while smaller farms face barriers to meaningful participation.

Applications in Collaborative Agriculture

Federated Yield Prediction and Crop Modeling

Crop yield prediction represents a fundamental application where federated learning enables farms to collaboratively develop accurate forecasting models while protecting proprietary productivity information. Individual farms typically possess multi-year records of crop yields associated with weather conditions, soil characteristics, management practices, and genetic varieties, but limited geographic scope constrains model generalization. Federated learning allows farms across regions to jointly train yield prediction models that capture diverse environmental responses and management effects without sharing actual productivity data that could reveal competitive information.

Horizontal federated learning architectures are particularly suitable for yield prediction, where participating farms collect similar feature sets including remote sensing indices, meteorological variables, soil properties, and phenological observations but represent different instances. Local training on farm-specific historical data enables models to learn local environmental relationships, while secure aggregation synthesizes knowledge across geographic gradients and management systems. Federated models can achieve accuracy comparable to centralized approaches while maintaining privacy, with studies demonstrating effective yield forecasting for major crops using distributed farm data.

Disease Surveillance and Pest Forecasting Systems

Early detection and forecasting of crop diseases and pest outbreaks benefit substantially from collaborative multi-farm intelligence that can identify regional patterns and predict spread dynamics. However, individual farms may be reluctant to share disease occurrence data due to concerns about market perception, regulatory implications, or competitive disadvantages. Federated learning enables farms to contribute to regional disease surveillance systems while keeping specific incidence information confidential.

Distributed deep learning models trained federatively on image data from farm-deployed cameras, drone imagery, and handheld devices can detect disease symptoms, identify pest species, and quantify severity levels across multiple operations. Secure aggregation ensures that disease presence at individual farms remains private while enabling development of robust detection models trained on diverse symptom presentations and environmental contexts. Temporal federated learning frameworks can track disease progression and predict outbreak trajectories by synthesizing observations across farms without centralizing sensitive epidemiological data.

Cross-farm Benchmarking and Decision Support

Farmers benefit from understanding their performance relative to peers operating under similar conditions, but traditional benchmarking requires sharing detailed operational and financial data that many farms consider confidential. Federated analytics enable privacy-preserving benchmarking where farms can compare their practices and outcomes against aggregated regional statistics without revealing individual performance metrics. Secure multi-party computation protocols allow farms to jointly calculate percentile rankings, identify efficiency frontiers, and detect optimal practice combinations while learning only their relative position within the distribution.

Collaborative decision-support systems leveraging federated learning can provide personalized recommendations derived from collective agricultural knowledge. For example, federated recommendation systems might suggest optimal planting dates, variety selections, or nutrient management strategies by learning from successful outcomes across similar farms while adapting recommendations to individual farm characteristics. Privacy-preserving federated optimization can identify Pareto-optimal solutions balancing productivity, profitability, and environmental outcomes across multi-farm datasets without centralized data aggregation.

Challenges and Future Perspectives

Communication Efficiency and Network Infrastructure

Agricultural environments frequently feature limited internet connectivity, unstable network conditions, and restricted bandwidth that challenge federated learning implementations requiring frequent model update transmission. Communication costs can dominate computational costs in federated settings, particularly when training large deep learning models with millions of parameters. Strategies for enhancing communication efficiency include model compression techniques such as quantization and pruning that reduce update sizes, federated distillation approaches where farms transmit predictions rather than gradients, and hierarchical federation architectures that aggregate updates regionally before global aggregation.

Edge computing infrastructure positioned at farms or regional agricultural service centers can facilitate federated learning by providing local computational resources and serving as aggregation points for geographically clustered farms. Investment in rural broadband infrastructure and development of communication protocols optimized for intermittent connectivity will be essential for scaling federated agricultural intelligence beyond pilot deployments.

Data Heterogeneity and Model Convergence

Agricultural data exhibits substantial heterogeneity arising from diverse soil types, climatic zones, crop varieties, management philosophies, and measurement protocols employed across farms. This non-independent and non-identically distributed data characteristic challenges federated learning convergence, potentially causing models to oscillate or converge to suboptimal solutions. Personalized federated learning approaches that maintain both global shared knowledge and farm-specific model components can accommodate heterogeneity while preserving collaborative benefits.

Meta-learning frameworks enable federated systems to learn how to rapidly adapt global models to individual farm contexts, while multi-task federated learning can simultaneously optimize related but distinct objectives across farms. Robust aggregation algorithms that down-weight contributions from statistical outliers or detect and exclude adversarial participants improve convergence in heterogeneous agricultural settings.

Scalability and Computational Resource Constraints

Many agricultural operations, particularly smaller farms in developing regions, possess limited computational infrastructure that may constrain participation in resource-intensive federated learning processes. Asymmetric federated learning architectures that assign computationally demanding tasks to better-resourced participants while enabling lightweight contribution from resource-constrained farms can promote inclusive participation. Federated learning frameworks designed for mobile and embedded devices employed in precision agriculture show promise for enabling broader participation without requiring substantial computational investments.

Scalability to large numbers of participating farms introduces coordination challenges including participant selection strategies, aggregation algorithm complexity, and fault

tolerance mechanisms. Hierarchical and clustered federation architectures that organize farms into regional groups with intermediate aggregation layers can improve scalability while potentially enhancing model relevance through geographic stratification.

Standardization and Interoperability

The agricultural technology ecosystem encompasses diverse farm management platforms, sensor systems, data formats, and ontologies that complicate federated learning deployment across heterogeneous technology environments. Development of standardized data schemas, application programming interfaces, and federated learning protocols specific to agricultural domains would facilitate interoperability and reduce integration barriers. Industry consortia, agricultural standards organizations, and open-source federated learning platforms tailored for farming applications could accelerate adoption by reducing implementation complexity.

Interoperability between federated learning systems and existing agricultural decision support tools, farm management information systems, and precision agriculture platforms requires careful interface design and workflow integration. Modular federated learning frameworks that can plug into diverse agricultural software ecosystems will be essential for practical deployment.

Policy, Regulation, and Adoption Incentives

Regulatory frameworks governing agricultural data, privacy, and collaborative intelligence remain nascent in many jurisdictions, creating uncertainty regarding legal compliance for federated learning implementations. Clear guidelines addressing data protection requirements, liability for model errors, intellectual property rights in collaboratively-trained models, and antitrust considerations for farmer cooperation would provide needed regulatory clarity. Policymakers should consider how data governance regulations can enable privacy-preserving collaboration rather than imposing barriers to beneficial knowledge sharing.

Adoption of federated learning in agriculture requires demonstration of tangible benefits that justify participation costs including data preparation, system integration, and ongoing maintenance. Incentive structures might include preferential access to high-quality collaborative models, reduced insurance premiums for participants in risk monitoring networks, premium pricing for products from farms employing data-driven sustainability practices verified through federated systems, or public investment in federated infrastructure as agricultural extension services. Successful case studies demonstrating measurable improvements in productivity, profitability, or sustainability from federated learning participation will be crucial for building farmer trust and motivating adoption.

Conclusion

Federated learning represents a transformative paradigm for enabling secure multi-farm data sharing and collaborative intelligence in digital agriculture while preserving privacy, respecting data sovereignty, and addressing legitimate concerns about competitive sensitivity. By distributing machine learning training processes to individual farms and aggregating only model updates rather than raw data, federated architectures unlock the collective value of distributed agricultural datasets without requiring centralized data repositories that many stakeholders find unacceptable. Advanced privacy-preserving mechanisms including secure aggregation, differential privacy, homomorphic encryption, and blockchain-based trust management provide robust protections for sensitive farm information throughout collaborative learning processes.

Agricultural applications including federated yield prediction, distributed disease surveillance, cross-farm benchmarking, and collaborative decision support demonstrate substantial potential for improving farming outcomes through collective intelligence. However, realizing this potential at scale requires addressing significant technical challenges related to communication efficiency in rural networks, accommodating data heterogeneity across diverse farming systems, ensuring model convergence with non-identically distributed datasets, and developing governance frameworks that balance collaborative benefits with individual farm interests.

Future progress depends on interdisciplinary collaboration integrating distributed machine learning research, agricultural domain expertise, cybersecurity innovation, and stakeholder engagement to develop practical federated learning systems tailored for real-world farm environments. Standardization efforts, supportive policy frameworks, demonstration of tangible adoption benefits, and investment in rural digital infrastructure will be essential enablers for transitioning federated agricultural intelligence from promising concept to widespread practice. As precision agriculture continues expanding, federated learning offers a privacy-preserving pathway toward collaborative intelligence that can accelerate sustainable intensification, climate adaptation, and equitable access to data-driven decision support across the global agricultural sector.

Table 1: Types of agricultural data suitable for federated learning in multi-farm environments

Data Category	Examples	Federated Learning Suitability	Privacy Considerations
Crop performance data	Yield records, quality metrics, harvest timing	High - enables collaborative yield modeling	Commercially sensitive, competitive concerns
Environmental monitoring	Soil moisture, temperature, precipitation, solar radiation	High - supports regional environmental models	Generally low sensitivity, aggregatable
Management practices	Planting dates, variety selections, input applications	Medium - valuable for practice optimization	Moderate sensitivity, proprietary strategies
Disease and pest observations	Incidence records, severity assessments, treatment responses	High - critical for regional surveillance	Regulatory and market perception concerns
Remote sensing data	Satellite imagery, drone data, vegetation indices	Medium - complements ground observations	Variable depending on resolution and coverage
Phenological observations	Growth stages, flowering dates, maturity timing	High - improves crop modeling accuracy	Low sensitivity, scientifically valuable

Economic data	Input costs, market prices, profitability metrics	Low - extremely sensitive and heterogeneous	Very high sensitivity, limited sharing willingness

Table 2: Federated learning algorithms and aggregation strategies for agricultural applications

Algorithm/Strategy	Mechanism	Agricultural Application Suitability	Communication Efficiency	Heterogeneity Tolerance
Federated Averaging (FedAvg)	Weighted averaging of local model parameters	High - general purpose for most crop modeling tasks	Moderate - requires periodic synchronization	Low - assumes similar data distributions
Federated Stochastic Gradient Descent	Distributed gradient descent with central aggregation	Medium - suitable for convex agricultural optimization	Low - frequent communication rounds	Low - convergence challenges with heterogeneity
Personalized Federated Learning	Maintains global and farm-specific model components	Very High - accommodates diverse farm conditions	Moderate - similar to FedAvg	High - explicitly handles heterogeneity
Hierarchical Federated Learning	Regional aggregation before global synthesis	High - matches agricultural geographic structures	High - reduces long-distance communication	Moderate - regional clustering helps
Asynchronous Federated Learning	Accepts updates without synchronization requirements	High - accommodates variable farm connectivity	High - no waiting for slow participants	Moderate - weighted aggregation handles timing
Federated Meta-Learning	Learns adaptation strategies across farms	High - enables rapid local customization	Moderate - additional meta-learning overhead	Very High - designed for distribution shift

Table 3: Privacy, security, and trust mechanisms used in federated learning-based farm data sharing

Mechanism	Function	Privacy Protection Level	Computational Overhead	Agricultural Implementation Challenges
Secure Aggregation	Cryptographic protocols preventing individual update visibility	High - server learns only aggregates	Moderate - cryptographic operations required	Manageable for most farm systems
Differential Privacy	Noise injection providing statistical privacy guarantees	Very High - formal privacy bounds	Low to Moderate - noise computation	Balancing privacy-accuracy trade-offs for agricultural predictions
Homomorphic Encryption	Computation on encrypted model updates	Very High - updates never decrypted	Very High - computationally intensive	Significant challenge for resource-constrained farms
Blockchain-based Trust	Distributed ledger for transparent governance and auditing	Moderate - transparency not privacy	Moderate - blockchain consensus costs	Network infrastructure and energy requirements
Secure Multi-party Computation	Joint computation without revealing inputs	Very High - cryptographic security	High - multiple communication rounds	Complexity of implementation and coordination
Trusted Execution Environments	Hardware-based isolation for secure aggregation	High - hardware-enforced protection	Low - leverages existing hardware features	Limited availability in agricultural edge devices
Access Control and Authentication	Identity verification and authorization management	Moderate - controls who participates	Low - standard authentication protocols	Establishing identity management for farm networks

Table 4: Advantages, limitations, and implementation challenges of federated learning in multi-farm agricultural systems

Aspect	Advantages	Limitations	Implementation Challenges
Privacy and Data Sovereignty	Preserves farm data locality; maintains confidentiality; respects ownership; reduces regulatory compliance burden	Model updates may still leak information; requires additional privacy mechanisms; not completely risk-free	Selecting appropriate privacy-preserving techniques; balancing privacy and model quality; educating stakeholders
Model Performance	Accesses diverse data without centralization; improves generalization; captures regional variations; enables larger effective training sets	Heterogeneous data challenges convergence; communication constraints limit model complexity; coordination overhead	Managing non-IID data distributions; developing robust aggregation; optimizing communication-computation trade-offs
Collaboration Enablement	Facilitates knowledge sharing among competitors; reduces barriers to cooperation; enables collective intelligence; democratizes AI access	Requires trust in aggregation infrastructure; participation coordination complexity; free-rider problems	Establishing governance frameworks; creating participation incentives; ensuring equitable benefit distribution
Communication and Infrastructure	Reduces raw data transmission; enables edge computing utilization; accommodates intermittent connectivity	Rural network limitations; communication bottlenecks; requires synchronization; bandwidth consumption	Developing communication-efficient protocols; deploying edge infrastructure; optimizing model compression
Scalability and Resource Requirements	Distributes computational load; leverages existing farm systems; enables incremental participation	Coordination complexity increases with participants; resource constraints on small farms; aggregation computation grows	Creating scalable architectures; supporting heterogeneous computational capabilities; developing lightweight protocols
Standardization and Interoperability	Potential for industry-wide collaboration; reusable frameworks; knowledge transfer across crops/regions	Lack of agricultural federated learning standards; diverse technology ecosystems; integration complexity	Developing agricultural domain standards; ensuring cross-platform compatibility; building open-source tools

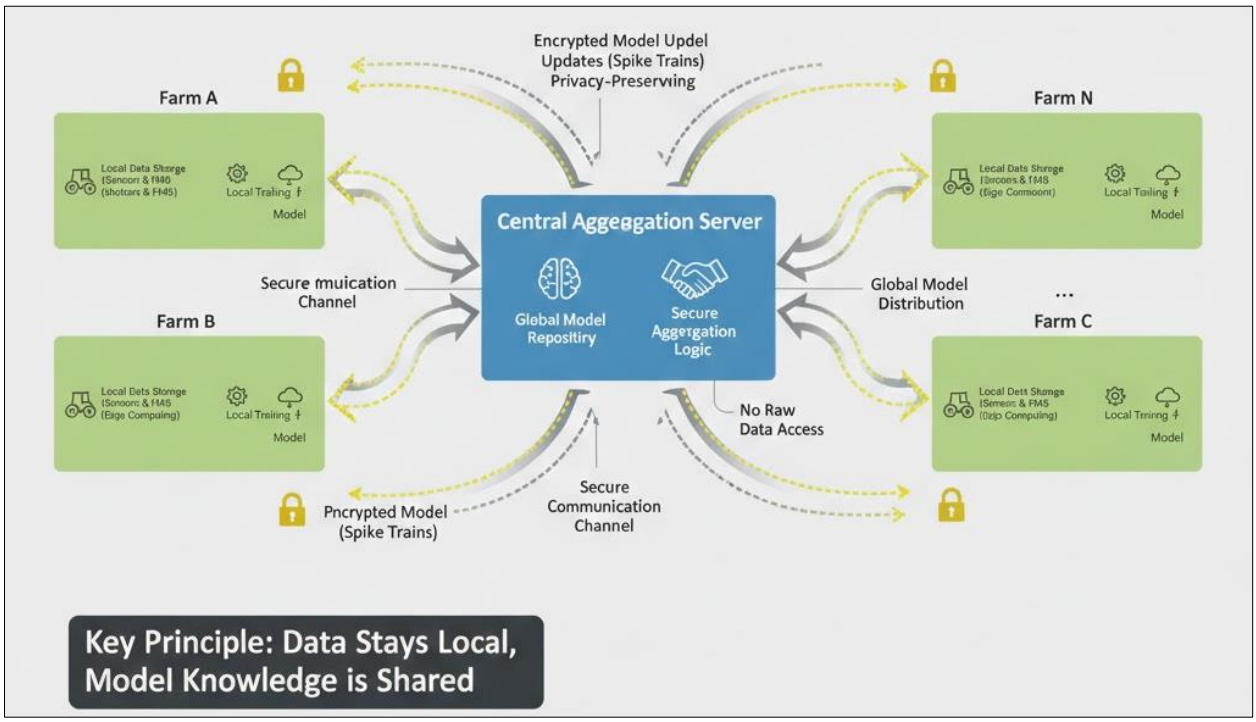

Fig 1: Conceptual architecture of federated learning for secure multi-farm data collaboration

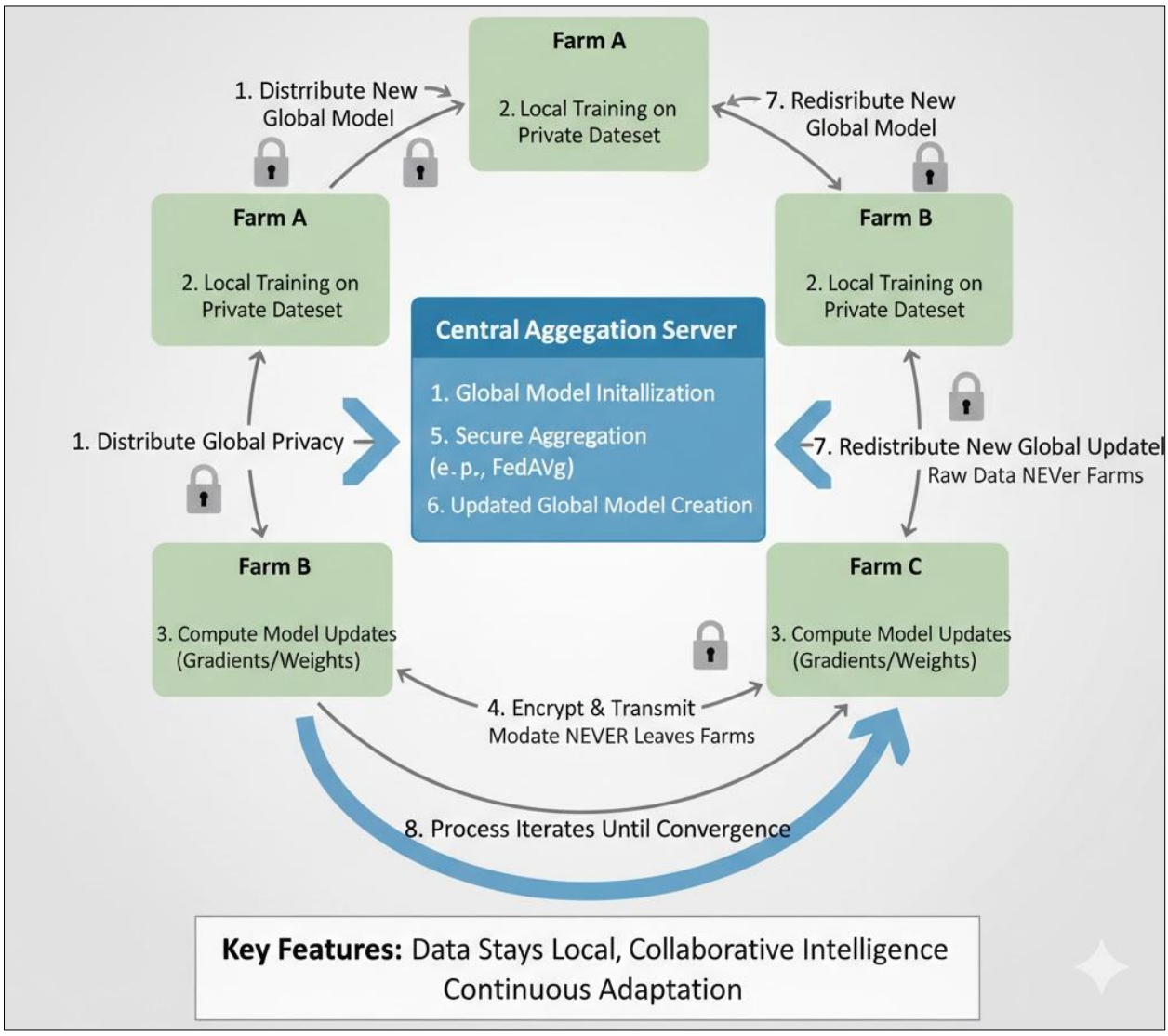

Fig 2: Federated learning workflow illustrating local training, secure aggregation, and global model updates across farms

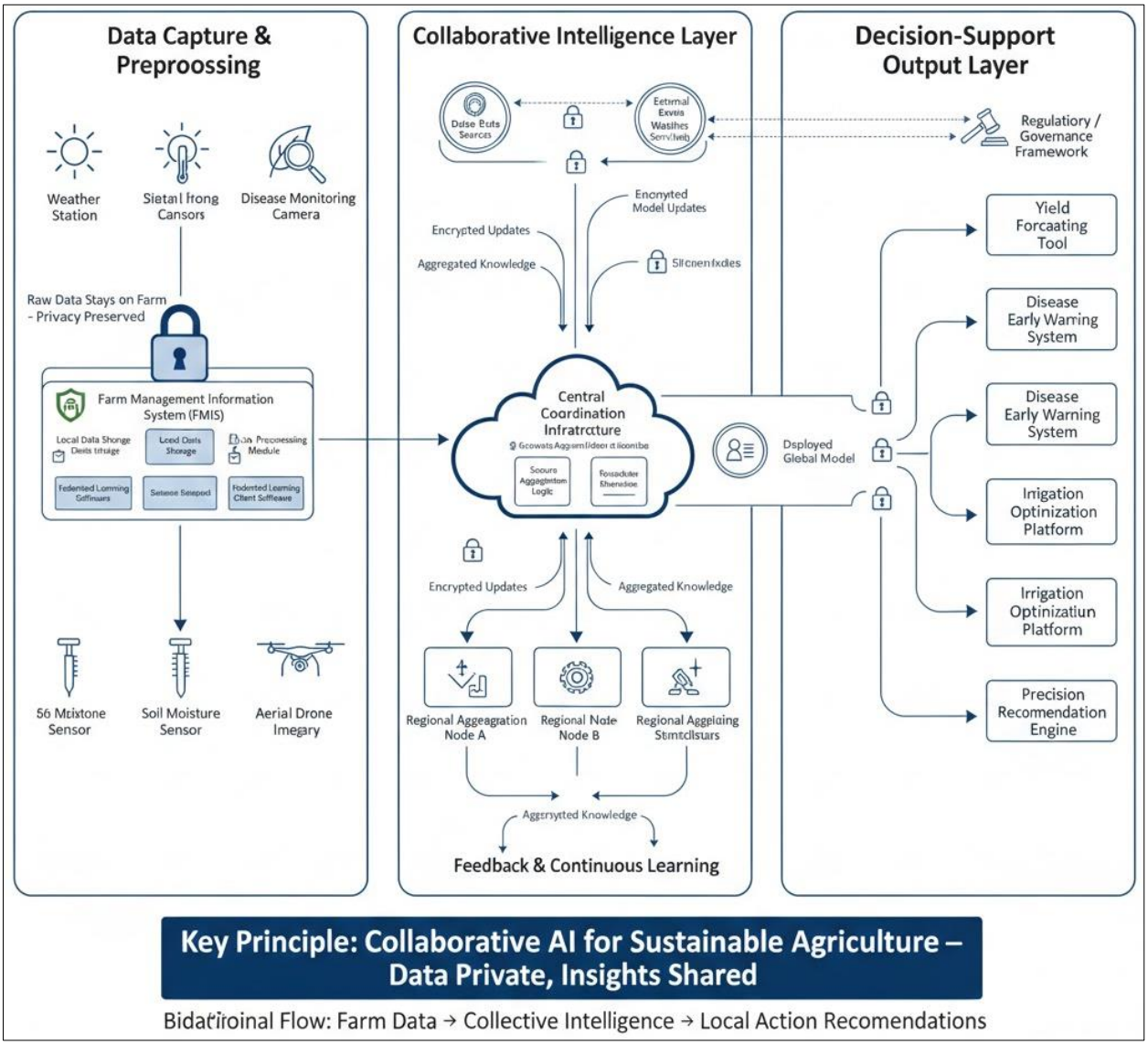

Fig 3: Integration of federated learning with agricultural sensors, farm management systems, and decision-support platforms

References

1. Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: A review. Sensors. 2018;18(8):2674.
2. Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture. 2018;147:70-90.
3. Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big data in smart farming – A review. Agricultural Systems. 2017;153:69-80.
4. Weersink A, Fraser E, Pannell D, Duncan E, Rotz S. Opportunities and challenges for big data in agricultural and environmental analysis. Annual Review of Resource Economics. 2018;10:19-37.
5. Kootstra G, Wang X, Blok PM, Hemming J, Van Henten E. Selective harvesting robotics: current research, trends, and future directions. Current Robotics Reports. 2021;2(1):95-104.
6. Zhai Z, Martínez JF, Beltran V, Martínez NL. Decision support systems for agriculture 4.0: Survey and challenges. Computers and Electronics in Agriculture. 2020;170:105256.
7. Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, *et al*. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural Systems. 2017;155:269-288.
8. Wiseman L, Sanderson J, Zhang A, Jakku E. Farmers and their data: An examination of farmers' reluctance to share their data through the lens of the laws impacting smart farming. NJAS - Wageningen Journal of Life Sciences. 2019;90-91:100301.
9. Carbonell IM. The ethics of big data in big agriculture. Internet Policy Review. 2016;5(1):1-13.
10. Jakku E, Taylor B, Fleming A, Mason C, Fielke S, Sounness C, *et al*. "If they don't tell us what they do with it, why would we trust them?" Trust, transparency and benefit-sharing in Smart Farming. NJAS - Wageningen Journal of Life Sciences. 2019;90-91:100285.
11. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology. 2019;10(2):1-19.
12. Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine. 2020;37(3):50-60.
13. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA. Communication-efficient learning of deep networks

from decentralized data. In: Artificial Intelligence and Statistics. PMLR; 2017. p. 1273-1282.
14. Konečný J, McMahan HB, Yu FX, Richtárik P, Suresh AT, Bacon D. Federated learning: Strategies for improving communication efficiency. arXiv preprint arXiv:1610.05492. 2016.
15. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, *et al*. Advances and open problems in federated learning. Foundations and Trends in Machine Learning. 2021;14(1-2):1-210.
16. Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, *et al*. A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering. 2021;35(4):3347-3366.
17. Aledhari M, Razzak R, Parizi RM, Saeed F. Federated learning: A survey on enabling technologies, protocols, and applications. IEEE Access. 2020;8:140699-140725.
18. Lim WYB, Luong NC, Hoang DT, Jiao Y, Liang YC, Yang Q, *et al*. Federated learning in mobile edge networks: A comprehensive survey. IEEE Communications Surveys & Tutorials. 2020;22(3):2031-2063.
19. Friha O, Ferrag MA, Shu L, Maglaras L, Wang X. Internet of things for the future of smart agriculture: A comprehensive survey of emerging technologies. IEEE/CAA Journal of Automatica Sinica. 2021;8(4):718-752.
20. Yang Q, Liu Y, Cheng Y, Kang Y, Chen T, Yu H. Federated learning. Synthesis Lectures on Artificial Intelligence and Machine Learning. 2019;13(3):1-207.
21. Chen Y, Qin X, Wang J, Yu C, Gao W. Fedhealth: A federated transfer learning framework for wearable healthcare. IEEE Intelligent Systems. 2020;35(4):83-93.
22. Liu Y, Kang Y, Xing C, Chen T, Yang Q. A secure federated transfer learning framework. IEEE Intelligent Systems. 2020;35(4):70-82.
23. Bonawitz K, Eichner H, Grieskamp W, Huba D, Ingerman A, Ivanov V, *et al*. Towards federated learning at scale: System design. In: Proceedings of Machine Learning and Systems. 2019. p. 374-388.
24. Wang H, Kaplan Z, Niu D, Li B. Optimizing federated learning on non-iid data with reinforcement learning. In: IEEE INFOCOM 2020. IEEE; 2020. p. 1698-1707.
25. Samarakoon S, Bennis M, Saad W, Debbah M. Distributed federated learning for ultra-reliable low-latency vehicular communications. IEEE Transactions on Communications. 2020;68(2):1146-1159.
26. Sattler F, Wiedemann S, Müller KR, Samek W. Robust and communication-efficient federated learning from non-iid data. IEEE Transactions on Neural Networks and Learning Systems. 2019;31(9):3400-3413.
27. Chen Y, Sun X, Jin Y. Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation. IEEE Transactions on Neural Networks and Learning Systems. 2020;31(10):4229-4238.
28. Zhang W, Wang X, Zhou P, Wu W, Zhang X. Client selection for federated learning with non-IID data in mobile edge computing. IEEE Access. 2021;9:24462-24474.
29. Vepakomma P, Gupta O, Swedish T, Raskar R. Split learning for health: Distributed deep learning without sharing raw patient data. arXiv preprint arXiv:1812.00564. 2018.
30. Zhu L, Liu Z, Han S. Deep leakage from gradients. In: Advances in Neural Information Processing Systems. 2019. p. 14774-14784.
31. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, *et al*. Practical secure aggregation for privacy-preserving machine learning. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017. p. 1175-1191.
32. Aono Y, Hayashi T, Wang L, Moriai S. Privacy-preserving deep learning via additively homomorphic encryption. IEEE Transactions on Information Forensics and Security. 2017;13(5):1333-1345.
33. Zhang C, Li S, Xia J, Wang W, Yan F, Liu Y. BatchCrypt: Efficient homomorphic encryption for cross-silo federated learning. In: 2020 USENIX Annual Technical Conference. 2020. p. 493-506.
34. Xu R, Baracaldo N, Zhou Y, Anwar A, Ludwig H. HybridAlpha: An efficient approach for privacy-preserving federated learning. In: Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security. 2019. p. 13-23.
35. Abadi M, Chu A, Goodfellow I, McMahan HB, Mironov I, Talwar K, *et al*. Deep learning with differential privacy. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. 2016. p. 308-318.
36. Geyer RC, Klein T, Nabi M. Differentially private federated learning: A client level perspective. arXiv preprint arXiv:1712.07557. 2017.
37. Wei K, Li J, Ding M, Ma C, Yang HH, Farokhi F, *et al*. Federated learning with differential privacy: Algorithms and performance analysis. IEEE Transactions on Information Forensics and Security. 2020;15:3454-3469.