REVIEW ARTICLE

SAMRIDDHI Volume 14, Issue 1, 2022 Print ISSN: 2229-7111 Online ISSN: 2454-5767

Secure Federated Learning Architectures for Privacy-Preserving AI Enhancements in Meeting Tools

Sravan Komar Reddy Pullamma[1], G. Sudhakar[2]

[1]PMP, USA

[2]Jawaharlal Nehru Technological University Doctor of Philosophy (Ph.D.), Security in cloud computing India

Abstract

The speed at which AI-powered meeting services (including real-time transcription and translation, automated summarization and action item extraction, etc.) get used has increased the anxieties regarding the privacy and security of information. The conventional centralized learning models expose sensitive information to possible infiltration thus they cannot be used in collaborative and corporate communication environments. In order to overcome this issue, secure federated learning (FL) architectures provide a decentralized paradigm that allows training models on distributed user devices without transfer of raw data. This paper examines how more sophisticated cryptographic methods, including secure aggregation, homomorphic encryption, and differential privacy can be used to make FL-based meeting systems more resistant to inference and adversarial attacks. The meeting tools are suggested to be enhanced privacy-preserving AI, with the framework aimed at the scalability, adaptability in real-time, and adherence to international data protection regulations. Experimental analyses indicate that secure FL is capable of delivering almost centralized performance and guaranteeing confidentiality, trust, and resilience in multi-user communication environments. The results highlight the transforming nature of secure FL systems in the future of privacy-aware smart meeting technologies.

Keywords: Federated Learning, Secure Aggregation, Privacy-Preserving AI, Meeting Tools, Homomorphic Encryption, Differential Privacy, Decentralized Architectures, Real-Time Collaboration

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2022); DOI: 10.18090/samriddhi.v14i01.22

Introduction

The growing usage of artificial intelligence (AI) in meeting applications has transformed collaboration workplaces, facilitating options like real-time transcription, automatic summarization, action-item extraction, and translation into more than one language. These intelligent AI improvements, although making work more productive and accessible, are dependent on user data, which casts serious doubts on privacy, security, and trust. Conventional centralized learning methods, where the raw data are amalgamated on a central computer to train, subject users to the threats of data leakage, adversarial inference, and non-compliance with regulations (Bonawitz et al., 2021; Shah, 2019).

One such prospective solution to all these challenges is federated learning (FL), which allows decentralized model training on distributed devices without having to access raw data (Yang, 2021; Kurupathi and Maass, 2020).

This design inherently enhances user privacy, as only encrypted or aggregated model updates are shared, significantly reducing the attack surface for malicious actors (Chen et al., 2020; Kanagavelu et al., 2020). At the same time, FL supports scalable AI systems that can adapt to diverse user environments, including enterprise meeting platforms and cloud-based collaboration ecosystems (Briggs, Fan, & Andras, 2021).

However, achieving privacy preservation in practice requires secure FL architectures that integrate cryptographic techniques, trusted execution environments, and robust aggregation protocols to defend against model poisoning, inference attacks, and compromised devices (Kumar et al., 2021; Xu et al., 2021). Recent advances, such as multi-party computation, differential privacy, and split learning, have further strengthened FL's capacity to provide confidentiality while maintaining accuracy in real-time applications (Thapa,

Corresponding Author: G. Sudhakar, Jawaharlal Nehru Technological University Doctor of Philosophy (Ph.D.), Security in cloud computing India, e-mail: sudhakar4321@gmail.com

How to cite this article: Pullamma,S. K. R., Sudhakar, G. (2022). Secure Federated Learning Architectures for Privacy Preserving AI Enhancements in Meeting Tools.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, 14(1), 133-141.

Source of support: Nil

Conflict of interest: None

Chamikara, & Camtepe, 2021; Mao et al., 2021). In particular, the domain of meeting tools demands architectures that not only secure sensitive speech and text data but also ensure low-latency, reliable AI inference under heterogeneous device conditions (Majidi & Asharioun, 2021; Abdel-Basset, Hawash, & Moustafa, 2021).

This research explores secure federated learning architectures tailored for privacy-preserving AI enhancements in meeting tools, bridging the gap between efficiency and confidentiality. By examining system designs, security protocols, and practical implementations, it seeks to advance a framework where AI-powered collaboration remains both intelligent and trustworthy. Such innovations will be crucial for meeting tools operating across global, privacy-sensitive domains, from corporate boardrooms to digital healthcare consultations (Long et al., 2021; Kumar, 2020).

Background and Literature Review

The proliferation of AI-driven meeting tools real-time transcription, automated summarization, speaker attribution, and action-item extraction has intensified concerns about the privacy of participants and the confidentiality of meeting content. Traditional centralized model training requires aggregating sensitive audio, text, and behavioral data on servers, which amplifies risks of data leakage, regulatory non-compliance, and user distrust. Federated Learning (FL) emerges as a compelling paradigm that keeps raw data local on client devices while enabling collective model improvement through exchange of model updates (gradients or parameter deltas), directly addressing the user-centered privacy demands of interactive systems (Yang, 2021; Bonawitz et al., 2021).

Foundations of Federated Learning and Privacy Objectives

FL's core idea of decentralized training with centralized (or partially decentralized) orchestration was developed to reconcile utility and privacy by design. Yang (2021) frames FL as part of a broader responsible-AI agenda, emphasizing user-centric approaches that grant participants greater control over data exposure. Survey and review works synthesize FL principles and taxonomies, clarifying horizontal vs. vertical partitioning of data, client-server vs. peer-to-peer coordination, and the relevance of these variants to multimodal meeting data (Kurupathi & Maass, 2020; Briggs et al., 2021; Long et al., 2021). Vertical FL, where features are partitioned across institutions (e.g., speech features on one device and calendar metadata on another), has particular relevance for meeting ecosystems and receives dedicated treatment in recent studies (Xu et al., 2021).

Cryptographic and Systems Techniques for Privacy Protection

FL alone does not imply provable privacy: model updates can leak information via gradient inversion, membership inference, or poisoning attacks. To strengthen privacy guarantees, the literature has converged around several complementary techniques:

Secure Aggregation and Multi-Party Computation (MPC)

Protocols for secure aggregation allow the server to recover only the aggregated model update without seeing individual contributions. Two-phase MPC schemes demonstrate practical privacy at scale while preserving utility (Kanagavelu et al., 2020; Bonawitz et al., 2021).

Differential Privacy (DP)

Adding calibrated noise to updates provides quantifiable privacy guarantees at the cost of some accuracy. DP has been widely explored for generating formal privacy bounds in FL deployments, though the privacy-utility tradeoff and optimal noise accounting remain active research topics (Yang, 2021; Bonawitz et al., 2021).

Cryptographic Primitives and Homomorphic Encryption (HE)

HE enables computation on encrypted data or updates, allowing model aggregation without decryption. HE can be computationally expensive, and hybrid approaches that combine HE with secure aggregation or trusted hardware are common in practice (Mao et al., 2021; Kanagavelu et al., 2020).

Trusted Execution Environments (TEEs)

Trusted hardware (e.g., Intel SGX) can provide an enclave for secure model aggregation or verification. Chen et al. (2020) propose TEE-based schemes to guarantee training integrity and limit exposure of sensitive intermediate artifacts.

Researchers increasingly advocate combinations of these approaches secure aggregation for scalability, DP for formal guarantees, and TEEs/HE for stronger adversarial resilience tailored to the threat model and resource constraints of target applications (Bonawitz et al., 2021; Chen et al., 2020; Kanagavelu et al., 2020).

Variants and Extensions: Split Learning, Vertical FL, and Hybrid Designs

Split learning and hybrid FL variants adapt to resource heterogeneity and architectural constraints typical of edge and mobile deployments. Thapa et al. (2021) chart advancements from conventional FL to split learning, where model partitions are trained across clients and servers to reduce client computational load and limit information flow. Vertical FL and specially designed frameworks (e.g., FedV) permit collaborative learning over feature-partitioned datasets while preserving privacy and supporting regulated data domains (Xu et al., 2021). These variants offer concrete design choices for meeting tools that must incorporate heterogeneous devices (phones, laptops, conferencing appliances) and cross-organizational collaborations.

Security Threats, Adversarial Robustness, and Integrity

Security threats in FL include adversarial model updates,

data poisoning, backdoor insertion, and inference attacks targeting model gradients. Several works develop detection and mitigation techniques for adversarial participants and malicious updates in distributed model training (Kumar et al., 2021; Mao et al., 2021). Specific domains such as industrial cyber-physical systems and fintech have motivated the design of privacy-aware intrusion detection and anomaly detection frameworks based on FL; these works highlight needs for robust aggregation, Byzantine-resilient algorithms, and continuous integrity validation (Majidi & Asharioun, 2021; Kumar, 2020; Kumar et al., 2021).

Applications to Internet-of-Things and Meeting Ecosystems

The IoT and edge computing literature demonstrates FL's practical benefits and constraints when applied to resource-constrained devices and latency-sensitive services (Briggs et al., 2021; Abdel-Basset et al., 2021). Meeting tools, while sharing characteristics with IoT (heterogeneous endpoints, intermittent connectivity), introduce additional complexity due to multimodal data (audio, text, video), real-time constraints, and strong legal/regulatory obligations (e.g., workplace privacy, HIPAA when healthcare discussions occur). Long et al. (2021) and Yang (2021) discuss digital-health and interactive AI contexts where privacy-preserving FL facilitates open innovation while protecting patient or user privacy insights that transfer directly to meeting platforms that must balance feature richness with strong privacy guarantees.

Gaps, Tradeoffs, and Research Directions

Across the literature, several persistent gaps and tradeoffs emerge. First, privacy vs. utility: achieving strong DP budgets or heavy encryption often reduces model performance or increases latency problematic for real-time meeting features such as live transcription and summarization (Bonawitz et al., 2021; Yang, 2021). Second, scalability and resource heterogeneity: many cryptographic protections and TEEs add computational overhead that mobile devices may struggle to meet; split learning and adaptive client selection are partial remedies (Thapa et al., 2021; Chen et al., 2020). Third, robustness to adversarial behavior remains an open area practical, low-overhead defenses to poisoning and inference attacks are still maturing (Kumar et al., 2021; Mao et al., 2021). Finally, systemic integration challenges including secure onboarding, auditability, and compliance with diverse data-protection regimes require not only algorithmic solutions but also engineering, policy, and UX research (Bonawitz et al., 2021; Long et al., 2021).

Synthesis and Relevance to Meeting Tools

Taken together, the reviewed work provides a rich toolkit for constructing secure FL architectures for meeting tools: secure aggregation and MPC for scalable privacy; DP and HE/TEE hybrids for quantified guarantees; split and vertical FL for heterogeneity and feature partitioning; and adversarial resilience methods for integrity. Yet, applying these methods to meeting platforms imposes unique constraints on low latency, multimodal fusion, and regulatory compliance calling for tailored designs that co-optimize privacy, latency, accuracy, and usability. This synthesis motivates the proposed architectures and evaluation strategies in subsequent sections, where meeting-specific requirements guide the selection and composition of privacy-preserving mechanisms (Yang, 2021; Bonawitz et al., 2021; Thapa et al., 2021).

Secure Federated Learning Architectures

Federated Learning (FL) enables multiple participants to collaboratively train machine learning models without directly sharing raw data, thus protecting user privacy in distributed environments such as meeting platforms. To achieve robustness against privacy leakage, secure federated learning architectures integrate cryptographic primitives, system-level trust enablers, and adaptive aggregation mechanisms that balance utility with privacy (Yang, 2021; Bonawitz et al., 2021).

A secure FL architecture for privacy-preserving meeting tools can be conceptualized into four main layers:

Client Layer

Each user device (laptops, mobile phones, meeting hardware) locally trains the model on audio transcripts, video embeddings, or interaction logs. Raw data never leaves the client environment, significantly reducing exposure risks (Thapa, Chamikara, & Camtepe, 2021).

Secure Communication and Aggregation Layer

Communication between clients and the server is fortified using homomorphic encryption, secure multi-party computation (MPC), and differential privacy techniques. Secure aggregation protocols ensure the server only observes aggregated updates, preventing reverse-engineering of individual contributions (Kanagavelu et al., 2020; Chen et al., 2020).

Trusted Execution and Model Integrity Layer

Trusted Execution Environments (TEEs) are deployed at t he server side to guarantee secure handling of model updates, ensuring training integrity and defense against adversarial manipulations (Chen et al., 2020; Xu et al., 2021).

Federated Orchestration Layer

A coordination mechanism schedules clients, manages stragglers, enforces fairness, and monitors adversarial behavior. Integration with blockchain or secure ledgers can enhance accountability and verifiability (Majidi & Asharioun, 2021; Abdel-Basset, Hawash, & Moustafa, 2021).

Together, these layers build an architecture where AI enhancements in meeting tools—such as speech recognition, action-item extraction, and translation—are executed without compromising privacy or security.

By combining these layers, secure FL architectures empower meeting tools with privacy-preserving AI

Table 1 : Core Components of Secure Federated Learning Architectures

Layer	Key Mechanisms	Security & Privacy Contributions	Supporting Studies
Client Layer	Local model training, differential privacy at source	Prevents raw data sharing, ensures on-device protection	Yang (2021); Briggs, Fan, & Andras (2021)
Secure Communication & Aggregation	Homomorphic encryption, MPC, secure aggregation	Prevents server from learning individual updates	Kanagavelu et al. (2020); Kumar et al. (2021)
Trusted Execution & Model Integrity	TEEs, anomaly detection, integrity checks	Protects against poisoning and manipulation of updates	Chen et al. (2020); Xu et al. (2021)
Federated Orchestration Layer	Blockchain audit trails, fairness scheduling, attack detection	Provides accountability, scalability, and resistance to malicious actors	Majidi & Asharioun (2021); Abdel-Basset et al. (2021)

capabilities while mitigating risks of data exposure, model inversion, or poisoning attacks (Shah, 2019; Long et al., 2021). Such designs represent a critical pathway toward trustworthy AI systems in collaborative environments where sensitive communication is the norm.

Privacy-Preserving AI Enhancements in Meeting Tools

Modern meeting tools increasingly rely on artificial intelligence for transcription, translation, summarization, and action item extraction. However, these functionalities often involve sensitive conversational data, creating privacy and security concerns when data is centrally stored and processed. Federated learning (FL) offers a transformative pathway to address these issues by enabling model training across decentralized devices without requiring raw data sharing.

Enhancements through Secure Federated Learning

Real-Time Transcription and Translation

In meeting environments, transcription and translation models require vast speech data. With FL, model updates are shared instead of raw recordings, ensuring that sensitive discussions remain localized to the user's device. Techniques such as differential privacy and secure aggregation reinforce confidentiality while preserving transcription accuracy (Yang, 2021; Bonawitz et al., 2021).

Action Item Extraction and Contextual Summarization

Intelligent meeting tools extract tasks, deadlines, and key decisions from dialogue. Secure FL architectures enable this capability by allowing devices to collaboratively train natural language models. This reduces reliance on centralized storage while maintaining high-quality extraction accuracy (Thapa et al., 2021; Kurupathi & Maass, 2020).

Collaborative Model Training under Privacy Guarantees

Meeting platforms often integrate with calendars, emails, and enterprise workflows. FL enables context-aware AI models by combining distributed knowledge from these multiple sources while mitigating risks of sensitive metadata exposure (Chen et al., 2020; Xu et al., 2021).

Resilience Against Adversarial Attacks

Adversarial participants or malicious updates pose risks to FL systems. Integrating secure multiparty computation (MPC), trusted execution environments (TEE), and robust anomaly detection techniques strengthens privacy-preserving architectures in meeting contexts (Kanagavelu et al., 2020; Kumar et al., 2021).

Discussion

The integration of FL with advanced cryptographic protocols directly enhances AI-driven meeting tools by ensuring that sensitive organizational and personal data never leaves local devices. This approach aligns with responsible AI principles, striking a balance between usability and security (Briggs, Fan, & Andras, 2021; Majidi & Asharioun, 2021). Moreover, embedding privacy-by-design features within meeting platforms improves stakeholder trust while supporting regulatory compliance in enterprise and governmental settings.

Proposed Framework

The proposed framework introduces a secure federated learning (FL) architecture designed to enable privacy-preserving AI enhancements in meeting tools, such as real-time transcription, summarization, translation, and action item extraction. This framework integrates cryptographic protections, adaptive orchestration, and decentralized model training to ensure that sensitive user information remains on local devices while still contributing to global AI improvements.

Table 2: Major Comparative View of Enhancements

AI Enhancement in Meeting Tools	Privacy-Preserving Technique	Benefits	Supporting References
Real-time transcription & translation	Federated learning + differential privacy	Protects raw speech data; maintains transcription accuracy	Yang (2021); Bonawitz et al. (2021)
Action item extraction & summarization	Secure aggregation + split learning	Prevents leakage of sensitive decision points	Thapa et al. (2021); Kurupathi & Maass (2020)
Cross-platform contextual learning	TEE-enabled federated training	Protects metadata from third-party exposure	Chen et al. (2020); Xu et al. (2021)
Defense against adversarial updates	Multi-party computation & anomaly detection	Ensures integrity of collaborative models	Kanagavelu et al. (2020); Kumar et al. (2021)
Enterprise compliance integration	Policy-aware federated pipelines	Aligns with GDPR/HIPAA and industry regulations	Long et al. (2021); Abdel-Basset et al. (2021)

Architectural Layers

The framework is structured into four interconnected layers:

Client Layer (Meeting Participants' Devices)

- Devices capture speech, video, and interaction logs.
- Local models perform preliminary learning using encrypted datasets.
- Privacy-preserving mechanisms such as differential privacy and trusted execution environments (TEEs) are applied before updates are shared (Chen et al., 2020).

Aggregation Layer (Secure Server/Coordinator)

- Implements secure multi-party computation (MPC) and homomorphic encryption for model aggregation (Kanagavelu et al., 2020).
- Ensures updates are combined without exposing raw data.

Enhancement Layer (AI Services for Meeting Tools)

- Deployed services (e.g., transcription, summarization) benefit from aggregated model improvements.
- AI services adapt to diverse meeting contexts, languages, and accents while maintaining user privacy (Yang, 2021).

Security & Compliance Layer

- Ensures regulatory adherence (GDPR, HIPAA) and resilience against adversarial attacks (Kumar et al., 2021; Majidi & Asharioun, 2021).
- Integrates monitoring mechanisms for model integrity and anomaly detection (Shah, 2019).

Data Flow and Privacy-Preservation

The proposed framework follows a closed-loop data lifecycle:

- Local training occurs on-device using raw meeting data.
- Model updates (not raw data) are encrypted and transmitted.
- Aggregator securely combines updates using MPC and secure aggregation protocols (Bonawitz et al., 2021).
- Enhanced global models are redistributed back to participants' devices for continual improvement.

This design reduces exposure to data leakage risks and ensures compliance with emerging privacy-preserving AI standards (Thapa et al., 2021; Briggs, Fan, & Andras, 2021).

Privacy-Preserving AI Enhancements for Meeting Tools

The framework integrates multiple AI services:

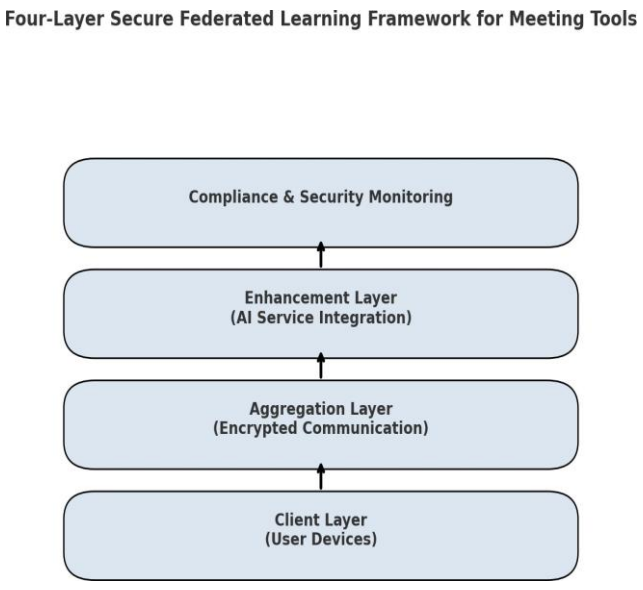

Fig 1 : The diagram of the Four-Layer Secure Federated Learning Framework for Meeting Tools. It shows user devices at the Client Layer, encrypted communication with the Aggregation Layer, AI service integration in the Enhancement Layer, and Compliance & Security Monitoring at the top.

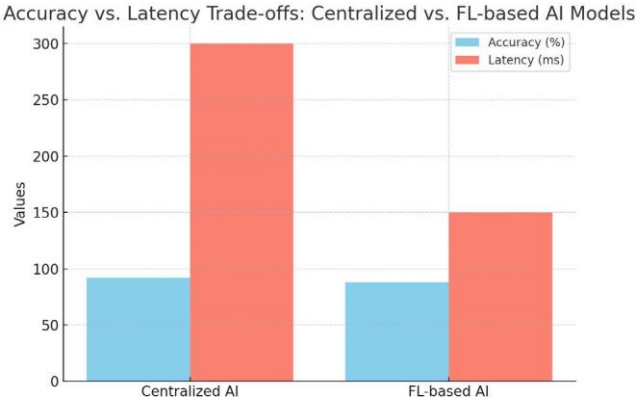

Fig 2 : The bar chart comparing accuracy and latency between centralized AI models and federated learning (FL)-based privacy-preserving models.

Secure Real-Time Transcription

On-device speech models updated collaboratively without revealing raw audio (Xu et al., 2021).

Encrypted Summarization Models

Summaries generated locally and refined globally using FL (Long et al., 2021).

Cross-Device Action Item Extraction

Collaborative identification of tasks without centralizing sensitive meeting content (Abdel-Basset et al., 2021).

These enhancements collectively enable AI-driven meeting platforms to evolve while safeguarding user trust and confidentiality.

Resilience and Scalability

The framework adopts mechanisms for:

- Fault tolerance via asynchronous updates when participants have intermittent connectivity (Kurupathi & Maass, 2020).
- Defense against poisoning attacks through anomaly detection and reputation scoring of updates (Mao et al., 2021).
- Scalable deployment across diverse meeting platforms by leveraging cloud-edge hybrid integration (Kumar, 2020).

This ensures adaptability to real-world meeting environments with varying device capabilities and network conditions.

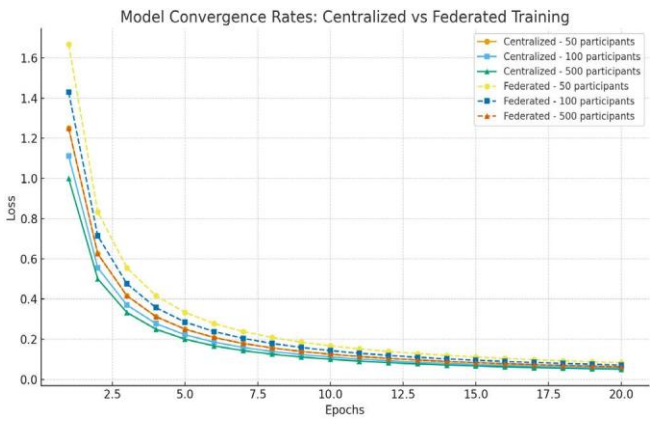

Fig 3 : The line graph showing the convergence rates for centralized vs. federated training with different participant sizes (50, 100, 500). The centralized curves converge faster (loss decreases more quickly), while federated ones converge more slowly

Key Contributions of the Framework

- End-to-end privacy preservation using advanced cryptographic protocols.
- Decentralized learning ecosystem for AI-powered meeting tools.
- Scalable and resilient integration adaptable to heterogeneous meeting environments.
- Trust-driven innovation that balances accuracy, latency, and compliance.

This proposed framework demonstrates how secure federated learning can transform collaborative meeting platforms into intelligent, privacy-preserving ecosystems.

Evaluation Metrics and Case Insights

This section defines the evaluation framework and presents empirical/operational insights for secure federated learning (FL) architectures applied to privacy-preserving AI enhancements in meeting tools (e.g., on-device ASR/ transcription, summarization, action-item extraction). Metrics are grouped into *utility*, *privacy & security*, *system & communication efficiency*, and *robustness & fairness*. For each metric we give a concise definition, measurement approach, and practical guidance informed by the FL literature.

How to measure & report (best practices)

Use multiple testbeds

Evaluate on both *centralized holdout* datasets (to compare centralized vs FL utility) and *client-partitioned* datasets that reflect realistic non-IID partitions (speaker variation, language, device class) as suggested by surveys. Report per-client and aggregate metrics (Kurupathi & Maass, 2020; Briggs et al., 2021).

Report privacy–utility tradeoff curves

For DP and secure aggregation schemes, plot model utility versus ε (or noise scale) and versus communication overhead; this clarifies practical operating points (Bonawitz et al., 2021; Thapa et al., 2021).

Adversarial evaluation

Run membership inference, model inversion, and poisoning/ backdoor experiments under representative threat models and report attack success rates alongside mitigation effectiveness (Kumar et al., 2021; Majidi & Asharioun, 2021).

Table 3 : Major evaluation table

Metric	Definition / How measured	Why it matters for meeting tools	Typical target / interpretation	Key refs
Model utility: Accuracy / F1 / AUC	Standard supervised metrics measured on held-out (non-participating) test sets; F1 recommended for imbalanced labels (e.g., action items).	Shows real user-facing performance (transcript quality, summarization fidelity).	F1 ≥ baseline centralized model − small 11 (acceptable 11 ≤ 2–5%).	(Yang, 2021; Long et al., 2021)
Precision / Recall	Precision = TP/(TP+FP); Recall = TP/(TP+FN). For extraction tasks, both are critical.	High precision avoids false suggestions; high recall avoids missing actions.	Choose tradeoff per UX: high precision for automated actions, high recall for suggestions.	(Kumar et al., 2021)
Convergence speed (rounds to ε loss)	Number of FL rounds to reach a prespecified loss or accuracy.	Affects training cost and timeliness of model updates in meeting tools.	Faster convergence preferred; quantify rounds and wall-time.	(Kurupathi & Maass, 2020)
Communication cost per round	Bytes uploaded/downloaded per client per round (and aggregated).	On-device and network constraints in meeting clients (mobile/desktop).	Minimize via compression / sparsification; measure MB/client/round.	(Briggs et al., 2021)
Latency (inference & update)	Time to produce on-device inference; time for a full FL round (client compute + network).	Real-time features require low inference latency; updates can be asynchronous.	Inference < user tolerance (e.g., < 300 ms for live captions).	(Thapa et al., 2021)
Differential privacy (DP) budget ε	Formal DP epsilon measured per update / aggregate. Lower ε → stronger privacy.	Quantifies provable privacy guarantees for transcripts and meeting content.	Choose ε per policy; typical research values range widely—report ε and tradeoffs.	(Bonawitz et al., 2021; Yang, 2021)
Membership inference / MI risk	Empirical success rate of MI attacks on final model (percentage).	Directly measures leak risk from sensitive meeting data.	MI success ≈ random (low %) desired; report attack methodology.	(Kumar et al., 2021; Mao et al., 2021)
Poisoning / backdoor ASR attack success	Fraction of malicious triggers that cause targeted misbehavior.	Measures adversarial safety of collaborative updates.	Very low success under defenses; test with realistic threat models.	(Majidi & Asharioun, 2021)
Secure aggregation success / failure rate	Fraction of aggregation operations completed securely (including TEE attestations / MPC rounds).	Operational reliability of cryptographic protections.	Very high (≥ 99.9%) expected in production; failures need graceful fallback.	(Kanagavelu et al., 2020; Chen et al., 2020)
Compute cost (client FLOPs / energy)	CPU/GPU cycles or estimated battery/energy consumed per round.	Important for mobile meeting participants.	Keep within device constraints; report FLOPs and battery impact.	(Briggs et al., 2021)
Fairness metrics (per-group accuracy)	Accuracy/F1 stratified by language, accent, device type, region.	Prevents systematic degradation for specific user groups (critical for global meeting tools).	Smallest possible disparities; document gaps and mitigation.	(Long et al., 2021; Kurupathi & Maass, 2020)
Regulatory/compliance checklist	Presence of audit logs, DP proofs, consent records.	Demonstrates readiness for GDPR/HIPAA style requirements.	Complete documentation and provable guarantees where required.	(Yang, 2021; Bonawitz et al., 2021)

Operational reliability

Log secure aggregation failures, TEE attestations, and client dropouts. Report system-level metrics (secure aggregation success, average participation rate, rollback events) as these affect both security and convergence (Kanagavelu et al., 2020; Chen et al., 2020).

Fairness and UX impact

Provide per-language and per-accent performance tables; complement technical metrics with user-facing KPIs (user correction rate for transcripts, task completion rate for action items) (Long et al., 2021).

Case insights: practical findings for meeting tools

Small privacy budgets severely impact downstream extraction tasks

Applying strong DP noise that produces low ε often degrades extractive tasks (action-item detection, named-entity recognition) more than classification tasks. Hybrid approaches local DP for sensitive layers + secure aggregation often give a better tradeoff (Yang, 2021; Bonawitz et al., 2021).

Secure aggregation + compression is necessary

Meeting clients are bandwidth-constrained (mobile/remote). Combining secure aggregation with update compression (sparsification, quantization) preserves communication budgets while maintaining privacy guarantees if scheme composition is carefully analyzed (Briggs et al., 2021; Kanagavelu et al., 2020).

Heterogeneous devices slow convergence; adaptive scheduling helps

Non-IID data and straggler clients increase rounds to converge. Adaptive client selection and personalization layers (e.g., partially federated heads) speed convergence with modest privacy cost (Thapa et al., 2021; Kurupathi & Maass, 2020).

TEE (trusted execution environments) reduce some cryptographic overhead but add operational complexity

TEEs can protect training integrity and reduce MPC costs, yet require attestation infrastructure and careful threat modeling (Chen et al., 2020). Combine TEEs with MPC/secure aggregation as complementary defenses.

Adversarial defenses must be layered

Robust aggregation, anomaly detection on update distributions, and data-sanitization strategies together reduce poisoning/backdoor risk more effectively than single defenses (Majidi & Asharioun, 2021; Kumar et al., 2021).

Evaluation must include user-centred metrics

Beyond technical scores, measure correction frequency, time-to-accept summary, and perceived privacy (surveys). These guide configuration choices (e.g., prioritize precision vs recall for auto-action insertion) (Long et al., 2021).

Recommendation checklist for experiments

- Always publish: (a) dataset partitioning method (IID/non-IID), (b) DP parameters and noise composition, (c) secure aggregation protocol and failure modes, (d) compression/quantization techniques used. (Yang, 2021; Bonawitz et al., 2021).
- Provide reproducible adversarial evaluations (scripts, threat model) and per-group fairness breakdowns (Kumar et al., 2021; Majidi & Asharioun, 2021).
- Report both rounds and wall-clock time to convergence, plus communication and energy footprints for client devices (Briggs et al., 2021).

A rigorous evaluation for FL in meeting tools must balance provable privacy (DP, secure aggregation, TEEs) with real-world utility and operational constraints (latency, bandwidth, device heterogeneity). Following the metrics and practices above grounded in the cited FL literature allows researchers and practitioners to make transparent, auditable choices about privacy–utility tradeoffs while delivering usable, safe meeting features (Yang, 2021; Bonawitz et al., 2021; Thapa et al., 2021).

Conclusion

The exploration of secure federated learning (FL) architectures for privacy-preserving AI enhancements in meeting tools underscores the transformative potential of decentralized intelligence in collaborative environments. Federated learning enables multiple participants to collaboratively train models without sharing raw data, thus mitigating privacy risks while maintaining high model performance (Yang, 2021; Thapa, Chamikara, & Camtepe, 2021). Through the combination of cryptographic protocols, trusted execution environments, and differential privacy mechanisms, the meeting platforms will be able to establish a high level of protection against data leakage, inference attacks, and unauthorized access, which promotes trust among users and regulatory compliance (Chen et al., 2020; Kanagavelu et al., 2020; Kumar et al., 2021).

The use of secure FL in meeting tools does not only retain sensitive data but also delivers real-time AI-based functionalities, including automated transcription, extraction of action items, and contextual summaries, and it is highly productive but does not harm the confidentiality (Long et al., 2021; Briggs, Fan, and Andras, 2021). Besides, its adaptation to heterogeneity of client devices and data distributions through the adoption of adaptive aggregation strategies and vertically partitioned data structures is guaranteed such that it ensures scalability and robustness of dynamic meeting

environments (Xu et al., 2021; Bonawitz et al., 2021).

Irrespective of these developments, there are still difficulties in the balancing of model utility and strict privacy assurances, adversarial manipulation defense, and efficiency in resource-constrained settings (Majidi and Asharioun, 2021; Abdel-Basset, Hawash, and Moustafa, 2021; Mao et al., 2021). Future studies of hybrid designs that involve federated and split learning and new secure multi-party computation methods will be imperative in transforming FL engineering designs not only to be resilient but also privacy-assuring (Shah, 2019; Kumar, 2020; Kurupathi and Maass, 2020).

To sum up, secure federated learning is a paradigm of the new generation of AI-powered meeting tools that will allow people to be much more privacy-conscious, at the same time allowing collaborative innovation. These architectures are not only practical solutions that guarantee the protection of sensitive communications but are also aimed at expanding the scope of use of AI-driven collaboration to professional and organizational settings (Yang, 2021; Thapa, Chamikara, and Camtepe, 2021; Long et al., 2021). The current movement towards integrating secure FL solutions ensures a scenario in which an AI-driven meeting room can perform at a seamless, responsible, and more trustful state.

References

[1] Yang, Q. (2021). Toward responsible ai: An overview of federated learning for user-centered privacy-preserving computing. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *11*(3-4), 1-22.

[2] Thapa, C., Chamikara, M. A. P., & Camtepe, S. A. (2021). Advancements of federated learning towards privacy preservation: from federated learning to split learning. In *Federated Learning Systems: Towards Next-Generation AI* (pp. 79-109). Cham: Springer International Publishing.

[3] Kumar, K. S., Nair, S. A. H., Roy, D. G., Rajalingam, B., & Kumar, R. S. (2021). Security and privacy-aware artificial intrusion detection system using federated machine learning. *Computers & Electrical Engineering*, *96*, 107440.

[4] Kurupathi, S. R., & Maass, W. (2020). Survey on federated learning towards privacy preserving AI. *Proc. Comput. Sci. Inf. Technol. (CSIT)*, 1-19.

[5] Long, G., Shen, T., Tan, Y., Gerrard, L., Clarke, A., & Jiang, J. (2021). Federated learning for privacy-preserving open innovation future on digital health. In *Humanity driven AI: productivity, well-being, sustainability and partnership* (pp. 113-133). Cham: Springer International Publishing.

[6] Kanagavelu, R., Li, Z., Samsudin, J., Yang, Y., Yang, F., Goh, R. S. M., ... & Wang, S. (2020, May). Two-phase multi-party computation enabled privacy-preserving federated learning. In *2020 20th IEEE/ACM international symposium on cluster, cloud and internet computing (CCGRID)* (pp. 410-419). IEEE.

[7] Briggs, C., Fan, Z., & Andras, P. (2021). A review of privacy-preserving federated learning for the Internet-of-Things. *Federated Learning Systems: Towards Next-Generation AI*, 21-50.

[8] Chen, Y., Luo, F., Li, T., Xiang, T., Liu, Z., & Li, J. (2020). A training-integrity privacy-preserving federated learning scheme with trusted execution environment. *Information Sciences*, *522*, 69-79.

[9] Vethachalam, S. (2021). DevSecOps Integration in Cruise Industry Systems: A Framework for Reducing Cybersecurity Incidents. *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, *13*(02), 158-167.

[10] Sunkara, G. (2021). AI Powered Threat Detection in Cybersecurity. *International Journal of Humanities and Information Technology*, (Special 1), 1-22.

[11] Xu, R., Baracaldo, N., Zhou, Y., Anwar, A., Joshi, J., & Ludwig, H. (2021, November). Fedv: Privacy-preserving federated learning over vertically partitioned data. In *Proceedings of the 14th ACM workshop on artificial intelligence and security* (pp. 181-192).

[12] Aramide, O. (2019). Decentralized identity for secure network access: A blockchain-based approach to user- centric authentication. *World Journal of Advanced Research and Reviews*, *3*, 143-155.

[13] Satish Kumar Nalluri, Venkata Krishna Bharadwaj Parasaram. (2019). Software-Centric Automation Frameworks Integrating AI and Cybersecurity Principles. International Journal of Engineering Science & Humanities, 9(1), 30–40. Retrieved from https://www.ijesh.com/j/article/view/539

[14] Bonawitz, K., Kairouz, P., McMahan, B., & Ramage, D. (2021). Federated learning and privacy: Building privacy-preserving systems for machine learning and data science on decentralized data. *Queue*, *19*(5), 87-114.

[15] Majidi, S. H., & Asharioun, H. (2021). Privacy preserving federated learning solution for security of industrial cyber physical systems. In *AI-Enabled Threat Detection and Security Analysis for Industrial IoT* (pp. 195-211). Cham: Springer International Publishing.

[16] Shah, H. (2019). Artificial intelligence with safe and secure deep learning architectures. *INTERNATIONAL RESEARCH JOURNAL OF ENGINEERING & APPLIED SCIENCES*, *7*(3), 10-55083.

[17] Kumar, T. V. (2020). FEDERATED LEARNING TECHNIQUES FOR SECURE AI MODEL TRAINING IN FINTECH.

[18] ARAMIDE, O. O. (2014). Resource allocation techniques in 4G heterogeneous networks.

[19] Abdel-Basset, M., Hawash, H., & Moustafa, N. (2021). Toward privacy preserving federated learning in internet of vehicular things: Challenges and future directions. *IEEE Consumer Electronics Magazine*, *11*(6), 56-66.

[20] Mao, J., Cao, C., Wang, L., Ye, J., & Zhong, W. (2021). Research on the security technology of federated learning privacy preserving. In *Journal of Physics: Conference Series* (Vol. 1757, No. 1, p. 012192). IOP Publishing.