ARIC-FL: An Atomic Reputation-Aware Incentive Contract Framework for Blockchain-Enabled Federated Learning

Manar Ismail Abdelaziz
Faculty of Artificial Intelligence
Kafrelsheikh University, Egypt

Ahmed Ali Omar
Faculty of Artificial Intelligence
Kafrelsheikh University, Egypt

Rawan Shendy
Faculty of Artificial Intelligence
Kafrelsheikh University, Egypt

Nagham El-sayed
Faculty of Artificial Intelligence
Kafrelsheikh University, Egypt

Zainab H. Ali
Associate Professor at faculty of artificial intelligence kfs, School of applied science and engineering Nile University, Egypt

ABSTRACT
Federated learning (FL) enables distributed data owners to train a shared model without moving raw data to a central server. However, practical FL still faces an incentive and governance problem: clients spend computation, communication, energy, and privacy budget, while the task publisher cannot directly observe the true quality, reliability, or cost of each participant. Existing mechanisms based on Shapley value, contract theory, auctions, reinforcement learning, reputation, and blockchain smart contracts usually optimize separate parts of the incentive pipeline. This paper reviews representative incentive-driven and blockchain-enabled FL studies using six criteria: contribution evaluation, participant selection, reward allocation, privacy protection, attack resistance, and deployment cost. The comparative findings show that fair valuation remains computationally demanding, selection rules require independent contribution validation, reputation introduces cold-start bias, and blockchain improves auditability while adding transaction overhead. They also show that privacy-preserving valuation and atomic reward settlement are rarely combined with budget-aware participant selection in one lifecycle. Based on these findings, the paper proposes ARIC-FL, an Atomic Reputation-Aware Incentive Contract framework that combines reputation-aware reverse auction selection, adaptive contract menus, privacy-preserving contribution scoring, validator commit-reveal auditing, encrypted off-chain model exchange, and smart-contract escrow settlement. A review-based evidence matrix and gap-to-design traceability assessment evaluate the framework at the design level, while a reproducible protocol is retained for future empirical validation under IID, non-IID, and adversarial settings.

General Terms
Federated Learning, Blockchain, Security, Incentive Mechanisms, Algorithms.

Keywords
Federated learning; incentive mechanism; blockchain; smart contract; reputation; reverse auction; contract theory; Shapley value; secure aggregation; privacy-preserving machine learning.

1. INTRODUCTION
Federated learning (FL) is a major approach for collaborative machine learning when data are distributed across many owners and cannot be centralized. The model is moved toward the data, selected clients train locally, and only updates or parameters are returned for aggregation. This design reduces direct exposure of raw records and is useful in healthcare, finance, mobile services, Internet of Things (IoT), smart cities, and industrial edge systems [1], [2].

Nevertheless, FL privacy does not automatically imply FL reliability. Clients may be rational, unreliable, unavailable, or malicious. They invest computation, communication, energy, and privacy resources and may participate only if the expected reward is worth the cost. Poor incentive design can encourage free-riding, low-quality updates, collusion, poisoning, and payment disputes. Therefore, FL should be viewed not only as an aggregation problem, but also as a participation-management and accountability problem [3], [4].

Recent literature has proposed Shapley-value rewards, contract-theoretic pricing, budgeted auctions, reinforcement learning, reputation systems, and blockchain settlement. These directions are valuable but often fragmented. A fair valuation method can be computationally expensive; a transparent on-chain method can leak metadata; and an adaptive incentive rule can be difficult to audit. This paper therefore develops a pipeline view of incentive-driven FL in which recruitment, selection, local training, validation, aggregation, reward settlement, reputation update, and dispute handling are connected within a single framework [5], [6].

The main contribution of this paper is the proposal of ARIC-FL, an Atomic Reputation-Aware Incentive Contract framework for blockchain-enabled federated learning. ARIC-FL integrates reputation-aware reverse auction selection, adaptive contract menus, privacy-preserving contribution scoring, validator commit-reveal auditing, encrypted off-chain model exchange, and smart-contract escrow settlement. The framework aims to improve fairness, robustness, privacy protection, and reward reliability by ensuring that participants are selected based on their expected quality and cost, evaluated according to their validated contribution, and rewarded through an auditable atomic payment mechanism.

The remainder of this paper is organized as follows. Section 2, “BACKGROUND AND PRELIMINARIES,” presents the fundamental concepts of federated learning, incentive mechanisms, blockchain, smart contracts, and the associated

privacy and security challenges. Section 3, "REVIEW METHOD AND RELATED WORK," explains the review methodology and summarizes the representative incentive-driven and blockchain-enabled federated learning studies. Section 4, "COMPARATIVE ANALYSIS OF EXISTING METHODS," compares the reviewed approaches in terms of contribution evaluation, participant selection, reward allocation, privacy protection, attack resistance, and deployment overhead, and reports the main cross-study findings. Section 5, "RESEARCH GAP AND DESIGN REQUIREMENTS," identifies the limitations of the existing literature and derives the design requirements for the proposed framework. Section 6, "PROPOSED ARIC-FL FRAMEWORK," presents the architecture and main components of ARIC-FL, including reputation-aware selection, adaptive contracts, contribution scoring, validator auditing, atomic settlement, and reputation updating. Section 7, "SECURITY, PRIVACY, AND INCENTIVE ANALYSIS," analyzes the framework with respect to malicious clients, free-riding, payment default, validator collusion, privacy leakage, and incentive fairness. Section 8, "REVIEW-BASED EVALUATION AND FUTURE EMPIRICAL VALIDATION," provides a review-based assessment of the proposed framework and outlines the datasets, baselines, metrics, and adversarial scenarios recommended for future empirical validation. Section 9, "DISCUSSION AND LIMITATIONS," discusses the practical trade-offs, implementation challenges, and limitations of ARIC-FL. Section 10, "ETHICAL AND REGULATORY CONSIDERATIONS," addresses transparency, data privacy, reputation fairness, and responsible deployment. Finally, Section 11, "CONCLUSION AND FUTURE WORK," summarizes the findings and presents the main directions for future research.

2. BACKGROUND AND PRELIMINARIES

2.1 Federated Learning Model

In horizontal FL, a model owner coordinates a set of clients N = {1,...,n}. Client i owns a private dataset D_i with n_i samples. The global objective can be written as: min_w F(w) = sum_i (n_i/n)F_i(w), where F_i(w) is the local empirical loss of client i and n = sum_i n_i. This formulation follows the standard federated optimization setting used in FedAvg-based FL systems [1], [2]. A common aggregation step weights updates by data size, but data quantity is not always a sufficient proxy for value; a small client may contain rare, clean, or high-impact data. Therefore, incentive-aware FL should consider model utility, update quality, and participant reliability in addition to data volume [5], [7], [4].

2.2 Incentive Mechanisms in FL

An incentive mechanism determines how clients are selected, how contributions are measured, and how rewards are allocated. These three sub-problems (contribution evaluation, node selection, and payment allocation) are commonly identified as core components of incentive mechanism design in FL [4]. In a rational setting, a client participates only if p_i - c_i >= 0, where p_i is payment and c_i is participation cost. Useful mechanisms should be individually rational, incentive-compatible, budget-aware, fair, and resistant to manipulation [3], [4]. Contribution can be estimated using data volume, validation improvement, Shapley value, gradient similarity, reliability history, or hybrid scores [5], [7], [4].

2.3 Blockchain and Smart Contracts

Blockchain is best used as a coordination and audit layer, not as a storage layer for large model updates. Practical blockchain-enabled FL designs store task rules, bids, commitments, hashes, reputation states, and payment events on-chain, while encrypted model artifacts are exchanged off-chain using IPFS or secure peer-to-peer storage [8], [6], [9]. Smart contracts can escrow budgets, enforce deadlines, record evidence, release rewards, and penalize protocol violations [8], [10], [9].

2.4 Security and Privacy Issues

FL does not remove all privacy risks because gradients and model updates can leak information. Incentives can also increase strategic behavior: malicious clients may send random updates, poison labels, flip gradients, inflate data volume, delay training, or collude with validators [3], [7], [11], [6]. A robust incentive design must combine economic rules with validation, attack detection, secure aggregation, encryption, and privacy-preserving accountability [5], [9].

3. REVIEW METHOD AND RELATED WORK

The review corpus consists of ten representative studies covering incentive-driven FL, blockchain-enabled FL, smart-contract coordination, contract theory, auction mechanisms, reputation systems, reinforcement learning, and privacy-preserving contribution evaluation. It includes three survey or taxonomy papers and seven system-oriented papers. The synthesis follows a structured comparative approach rather than a statistical meta-analysis because the studies use different datasets, threat models, evaluation objectives, and implementation assumptions. Each study was examined according to six questions: how contribution is measured, how participants are selected, how rewards are allocated, how privacy is protected, which attacks are considered, and what practical overhead is introduced. For the system-oriented studies, the evidence matrix reports the dimensions strongly covered by each design and the dimensions that remain open or only partially addressed.

Survey and taxonomy studies provide the conceptual foundation for this paper. Zeng et al. [4] classified FL incentive mechanisms into major technical directions such as Shapley value, Stackelberg games, auctions, contract theory, reinforcement learning, and blockchain. Qu et al. [6] reviewed blockchain-enabled FL and emphasized that decentralization, consensus, and incentive mechanisms can mitigate centralization, data falsification, and lack of participant motivation. Ali et al. [3] further showed that incentive mechanisms should be analyzed together with security risks, including poisoning attacks, free-riding, spoofing, and privacy leakage.

System-oriented studies demonstrate how blockchain and smart contracts can coordinate training, validation, and rewards. Behera et al. [8] proposed a reward-driven smart-contract approach that records federated contribution on blockchain. Zhang et al. [9] presented Refiner, which uses workers, validators, smart contracts, and off-chain storage to address payment default and malicious updates. Gao et al. [7] proposed FGFL, where contribution and reputation jointly determine worker rewards and penalties, while Ouyang et al. [10] combined reverse auctions with reputation incentives in a blockchain-based decentralized FL model.

Other studies focus on economic and adaptive incentive design. Xuan et al. [12] used contract theory to model the more realistic case of multiple task publishers and multiple worker nodes. Han et al. [13] combined SVM-RFE selection, deep Q-network incentives, and proof-of-utility consensus for blockchain-based FL. More recently, Chen et al. [5] proposed πFL, which combines MPC-based Shapley valuation, weighted aggregation, and blockchain contracts to support privacy-preserving contribution assessment and atomic reward settlement.

As shown in Table 1, the reviewed literature provides strong building blocks, but most studies still emphasize a subset of the

incentive pipeline rather than integrating participant selection, contribution validation, privacy preservation, auditability, and atomic reward settlement into one deployable framework.

Table 1. Summary of representative studies

Study	Core method	Main strength	Remaining gap
Zeng et al. [4]	Survey of FL incentives: Shapley, Stackelberg, auction, contract, RL and blockchain.	Clarifies the design space.	No integrated deployable pipeline.
Qu et al. [6]	Survey of blockchain-enabled FL.	Links decentralization, incentives, security and applications	Open challenges remain architectural.
Ali et al. [3]	Systematic review of FL incentives and security.	Connects incentives to poisoning, free-riding and privacy.	Implementation architecture remains open.
Behera et al. [8]	Ethereum smart contracts and reward-driven FL.	Automates contribution recording and payment logic.	Validator accountability is limited.
Zhang et al. [9] - Refiner	Workers, validators, committee audit, smart contracts and off-chain exchange.	Addresses payment default and malicious behavior.	Validation quality and scale remain challenging.
Gao et al. [7] - FGFL	Contribution and reputation with blockchain incentive management.	Rewards useful workers and punishes attackers.	Cold-start and calibration issues.
Ouyang et al. [10] - BCD-FL	Reverse auction and reputation in blockchain FL with IPFS.	Efficient, individually rational and budget-aware.	Needs stronger collusion/privacy handling.
Xuan et al. [12]	Contract theory for multiple task publishers and workers.	Handles information asymmetry.	No blockchain audit or atomic settlement.
Han et al. [13]	SVM-RFE selection, DQN incentives and proof-of-utility consensus.	Combines dynamic incentives and blockchain trust.	Learned policies need safety constraints.
Chen et al. [5] - πFL	MPC-based Shapley, weighted aggregation and blockchain contracts.	Private contribution valuation and atomic reward.	MPC/Shapley cost remains high.

4. COMPARATIVE ANALYSIS OF EXISTING METHODS

4.1 Contribution Evaluation

Contribution evaluation is central to reward fairness. Data-volume rewards are cheap and transparent but easy to inflate, while validation-based rewards link payment to model quality but depend on the representativeness of the validation set. Leave-one-out and Shapley-value methods offer a stronger fairness interpretation, yet exact Shapley computation is expensive in large-scale FL [5], [4]. A practical system should therefore support multiple valuation levels: lightweight validation and gradient checks for low-risk contexts, and privacy-preserving MPC-assisted valuation for sensitive deployments.

4.2 Participant Selection

Random selection wastes rounds when clients are low quality, unavailable, or malicious. Auction-based selection handles budget limits, reputation improves long-term reliability, contract theory handles hidden client costs, and reinforcement learning adapts to changing conditions [13], [10], [12]. However, a strong design should combine price, predicted contribution, reputation, availability, and exploration so that new honest clients are not permanently excluded.

4.3 Reward Allocation

Equal rewards are simple but ignore quality. Data-volume rewards are partial and manipulable. Shapley rewards are fairer but computationally expensive, critical-payment auctions can encourage truthful bidding, and reputation-based rewards discourage repeated misbehavior [5], [7], [9]. ARIC-FL adopts a hybrid model: a base reward for accepted participation and a variable reward proportional to validated contribution and reputation, subject to a fixed escrowed budget.

4.4 Blockchain, Auditability and Privacy

Blockchain improves credibility through escrow, task registration, event logging, and payment release, but full model updates should remain off-chain to reduce storage and transaction overhead [8], [14], [6], [15]. Validator committees reduce dependence on a central server, yet they introduce collusion risks; therefore, commit-reveal manifests, random validator selection, staking, and slashing can improve accountability [9]. Privacy should be preserved by storing only hashes, commitments, and reward-relevant metadata on-chain [16], [17].

Table 2 compares the main incentive and blockchain-related methods at the mechanism level and extracts the design implications used to construct the proposed ARIC-FL framework.

Table 2. Method-level comparison and design implications

Method	Strength	Limitation	Design implication
Data volume	Low cost and transparent.	Inflation and quality blindness.	Insufficient alone.
Validation gain	Connected to model quality.	Validation bias risk.	Useful score component.
Shapley value	Strong fairness meaning.	High computation.	Use sampled or audit mode.
Contract theory	Handles hidden costs/types.	Does not validate honesty.	Combine with audit.
Reverse auction	Budget-aware and potentially truthful.	Collusion/cold-start risk.	Add reputation and exploration.
Reputation	Long-term reliability.	Cold-start and bias.	Use decay and appeal.
Reinforcement learning	Dynamic adaptation.	Harder to explain	Use safety constraints.
Smart contracts	Transparent settlement.	Transaction cost/latency.	Keep on-chain data minimal.

4.5 Cross-Study Review Findings

The comparative analysis yields five main findings. First, no single reviewed system covers all six evaluation criteria with the same level of maturity. Second, mechanisms with the strongest fairness interpretation, especially Shapley- and MPC-based valuation, also introduce the highest computational burden. Third, auctions and contract theory improve budget control and handle hidden participant costs, but they do not independently verify whether a submitted update is useful or honest. Fourth, reputation and attack-detection mechanisms improve long-term reliability but create cold-start, calibration, and potential manipulation problems. Fifth, blockchain strengthens auditability and payment enforcement, yet practical systems must move model artifacts and expensive validation off-chain to control latency and storage cost.

Table 3. Review-based evaluation of the seven system-oriented studies.

Study	Review-based evaluation
Behera et al. [8]	**Covered:** smart-contract reward lifecycle, contribution recording, and protected round communication **Remaining need:** validator accountability and broader adversarial testing
Refiner [9]	**Covered:** validator audit, payment-default handling, malicious-worker filtering, and off-chain exchange **Remaining need:** validation-set dependence and large-scale transaction cost
FGFL [7]	**Covered:** contribution-reputation scoring, attack detection, and reward-punishment governance **Remaining need:** reputation cold-start, calibration, and manipulation resistance
BCD-FL [10]	**Covered:** reputation-aware reverse auction, budget balance, truthful selection, and IPFS storage **Remaining need:** private contribution valuation and collusion or adaptive-poisoning analysis
Xuan et al. [12]	**Covered:** multi-publisher contract design, information asymmetry, and worker self-selection **Remaining need:** blockchain audit, update-quality validation, and atomic settlement
Han et al. [13]	**Covered:** learning-based client selection, adaptive incentives, and proof-of-utility consensus **Remaining need:** policy explainability, safety constraints, and scalability
πFL [5]	**Covered:** MPC-Shapley valuation, private contribution assessment, and atomic reward settlement **Remaining need:** MPC/Shapley cost and broader adversarial governance

Taken together, the study-level and mechanism-level comparisons show that the reviewed approaches are complementary rather than interchangeable. The survey studies establish the broader incentive-design space, whereas the system-oriented studies address selected components such as contribution valuation, participant selection, reputation management, attack detection, privacy-preserving assessment, and smart-contract settlement. This synthesis supports the need for an integrated incentive pipeline in which these functions are coordinated rather than optimized independently.

5. RESEARCH GAP AND DESIGN REQUIREMENTS

The review shows six recurring gaps: fragmented pipelines, the tradeoff between fair valuation and computational cost, privacy leakage during evaluation, reputation cold-start, incomplete payment atomicity, and strategic adaptation to scoring rules. These gaps are consistent with prior surveys on incentive-aware and blockchain-enabled FL, which emphasize that fairness, security, privacy, and deployment overhead must be addressed together rather than independently [3], [6], [4]. A practical framework should therefore satisfy fairness, individual rationality, budget balance, contribution sensitivity, privacy-preserving validation, resistance to malicious updates, validator accountability, low on-chain storage, cold-start fairness, and modular deployment [5], [7], [10], [9].

6. PROPOSED ARIC-FL FRAMEWORK

ARIC-FL stands for Atomic Reputation-Aware Incentive Contract Federated Learning. It is a modular framework for cross-device or cross-silo FL. The central principle is that reward should follow validated contribution, while validation and payment must remain private, auditable, and enforceable. The framework builds on reverse auction and reputation-based selection [10], contract-theoretic handling of hidden client information [18], [12], validator-based audit mechanisms [9], and privacy-preserving atomic reward settlement ideas [5].

6.1 System Model and Threat Model

ARIC-FL includes five roles: the task publisher defines the FL task and locks the budget; clients train locally and submit encrypted updates; validators evaluate updates using commit-

reveal; the smart contract records rules, commitments, scores, and payments; and off-chain storage holds encrypted model artifacts. Similar role separation and validator-based validation have been used in blockchain-enabled FL systems such as Refiner [9]. The framework considers rational clients, malicious clients, dishonest publishers, and colluding validators. It does not assume that blockchain alone prevents attacks; instead, it combines escrow, auditing, reputation, random committees, and penalties [3], [6], [9].

6.2 Reputation-Aware Reverse Auction

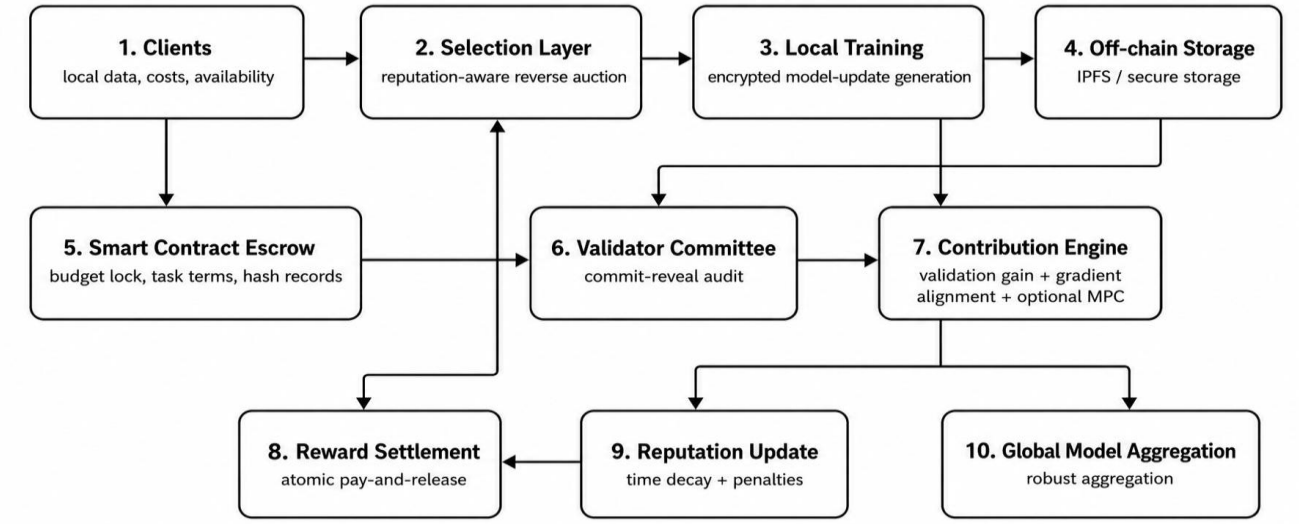

Fig 1: Proposed ARIC-FL architecture and incentive pipeline.

6.3 Adaptive Contract Menus

The publisher may publish contract options specifying privacy level, validation intensity, local epochs, deadlines, reward range, and penalty conditions. Clients self-select options that match their costs and privacy preferences. This follows the self-revealing logic of contract-theoretic FL incentive mechanisms, where hidden participant types and costs are addressed through contract menus [18], [12]. Thus, contract-theoretic ideas become executable smart-contract conditions rather than informal promises.

6.4 Multi-Factor Contribution Scoring

Accepted updates are scored using $Q_i^t = \alpha\Delta Acc_i^t + \beta Align_i^t + \gamma Novel_i^t + \delta Avail_i^t + \zeta PrivCost_i^t - \lambda Risk_i^t$. Validation improvement captures utility, alignment detects harmful gradients, novelty rewards underrepresented contribution, availability rewards timely delivery, privacy cost compensates stronger privacy requirements, and risk penalizes suspicious behavior. The score combines ideas from reputation-contribution assessment [7], Shapley/MPC-based valuation [5], and incentive taxonomy studies that emphasize contribution evaluation as a core FL incentive component [4]. The coefficients are declared before training.

6.5 Validator Committee and Atomic Settlement

Validators submit a commitment hash of the evaluation manifest, then reveal the manifest and salt after the commit phase closes. This reduces copying and late manipulation. The contract stores only hashes and final reward-relevant values. Atomic settlement is enforced by escrow: the publisher locks the round budget before training, rewards are released only after accepted validation evidence, and final model access is coupled with payment fulfillment. This design is aligned with validator-audited FL [9], auditable blockchain-based incentive systems [17], and atomic reward mechanisms for privacy-preserving FL [5].

Selection

At each round, clients submit bids representing required compensation. The selection score combines predicted contribution, reputation, availability, privacy cost, and bid price: $SelectScore_i^t = (E[Q_i^t] \cdot R_i^t \cdot A_i^t + \kappa \cdot Explore_i^t)/(b_i^t + \varepsilon)$. This design is motivated by reverse-auction incentive mechanisms [19], [10] and learning-based adaptive incentives [20]. The exploration term gives controlled opportunity to new or low-history clients, reducing reputation lock-in.

6.6 Reputation Update and Workflow

Reputation is updated after settlement using $R_i^{t+1} = \rho R_i^t + (1-\rho)ConductScore_i^t$, with time decay so old behavior matters less. New clients start with neutral reputation and can enter through exploration. This design follows reputation-based incentive governance in FL [7] and reputation-contract approaches for reliable FL participation [18]. The round workflow is: contract deployment, bidding, selection, local training, encrypted off-chain upload, validator commit-reveal, scoring, robust aggregation, reward release, reputation update, and next-round continuation.

7. SECURITY, PRIVACY AND INCENTIVE ANALYSIS

Individual rationality is supported because clients bid after observing contract terms and accepted honest clients receive base compensation plus a performance-sensitive reward. Budget balance is enforced because the smart contract cannot distribute more than the escrowed budget. Truthfulness is approached through auction pricing, self-selection, and contribution validation, following the desirable incentive properties discussed in auction-, contract-, and reverse-auction-based FL mechanisms [19], [10], [12], although a full implementation should formally prove the selected auction rule.

Malicious clients are discouraged by validation improvement, alignment checks, risk penalties, slashing, and reputation decay. Validator accountability is improved using random committees, role separation, and commit-reveal manifests [9]. Privacy is protected by keeping raw data local, encrypting off-chain

updates, storing only hashes and metadata on-chain, and enabling optional secure aggregation or MPC-assisted scoring in sensitive deployments [5], [16], [17]. Atomicity reduces publisher payment default and prevents clients from being paid without accepted evidence [5], [9].

8. REVIEW-BASED EVALUATION AND FUTURE EMPIRICAL VALIDATION

8.1 Design-Level Traceability Assessment

Because this paper is a comparative review and framework proposal, the present evaluation is design-level rather than experimental. The assessment traces each recurring gap identified in the reviewed literature to an explicit ARIC-FL component. It does not claim empirical superiority over the reviewed systems; instead, it verifies that the proposed architecture contains a concrete response to each requirement extracted from the comparison.

Table 4. Traceability from review findings to ARIC-FL design responses.

Review-derived requirement	ARIC-FL design response
Integrated incentive lifecycle	Connects recruitment, selection, training, validation, aggregation, settlement, reputation update, and dispute evidence in one workflow.
Fairness-cost trade-off	Uses lightweight validation and gradient checks by default, with optional MPC-assisted valuation for higher-risk deployments.
Privacy during evaluation	Keeps raw data local, encrypts off-chain updates, stores only commitments on-chain, and supports secure aggregation or MPC.
Cold-start and reputation bias	Adds an exploration term, neutral initial reputation, time decay, and appeal or audit evidence.
Atomic and budget-safe payment	Requires pre-funded escrow and couples accepted validation evidence with reward release and final model access.
Validator accountability and overhead	Uses random committees, commit-reveal, staking or slashing, and minimal on-chain metadata.

The traceability assessment shows why ARIC-FL is presented as an integrated enhancement rather than as an experimentally proven replacement for the reviewed methods. Empirical performance, scalability, and incentive compatibility remain subjects for the future protocol below.

8.2 Future Empirical Validation Protocol

ARIC-FL remains a conceptual and methodological framework; therefore, empirical performance claims require implementation and controlled experiments. Future validation should use MNIST or Fashion-MNIST for comparability, CIFAR-10 for a more difficult image task, and one application-specific dataset such as healthcare diagnosis, IoT anomaly detection, or financial fraud detection. Data should be partitioned under both IID and non-IID settings. Recommended baselines include FedAvg without incentives [1], equal reward allocation, data-volume reward, approximate Shapley reward, Refiner-style validation reward [9], FGFL-style contribution-reputation reward [7], BCD-FL-style reputation reverse auction [10], learning-based incentives [20], and πFL-style privacy-preserving contribution valuation [5]. Attack scenarios should include label flipping, sign flipping, Gaussian noise, free-riding, delayed updates, data-volume inflation, and validator collusion.

Table 5 lists the recommended empirical metrics for future validation, including model quality, incentive fairness, economic utility, security, privacy and cost, and reliability.

Table 5. Recommended empirical metrics for future validation.

Group	Concrete measures
Model quality	Accuracy, F1-score, test loss, convergence rounds and robustness under non-IID partitions.
Incentive fairness	Correlation between validated contribution and payment, Gini coefficient and reward share for high-quality clients.
Economic utility	Client utility, publisher utility, social welfare, budget utilization and budget balance.
Security	Attack success rate, poisoning resistance, free-rider detection, malicious reward leakage and validator disagreement.
Privacy/cost	Communication, computation, transaction count, off-chain storage and privacy-leakage proxy.
Reliability	Dropout tolerance, delayed update handling, cold-start selection rate and round completion time.

Ablation should remove one component at a time: reputation, exploration, validator commit-reveal, privacy-cost term, gradient-alignment term and atomic settlement. Results should be reported separately for clean IID, clean non-IID, adversarial, economic and overhead settings.

9. DISCUSSION AND LIMITATIONS

ARIC-FL is modular. A low-risk consortium may use escrow, reputation-based selection and encrypted off-chain exchange without MPC, while a healthcare deployment may activate stronger privacy-preserving scoring. The main tradeoff is trust versus cost: more validation, cryptography and auditing increase confidence but also increase overhead. The framework must therefore publish validation and privacy levels before clients bid. The main limitations are calibration and strategic adaptation. If validation gain is weighted too strongly, rare data may be undervalued; if reputation is weighted too strongly, new clients may be excluded; and if privacy cost is over-weighted, payments may increase without model improvement. Smart contracts enforce rules but cannot guarantee that the validation set is unbiased. Future implementations should include adversarial testing, periodic audit and appeal procedures.

10. ETHICAL AND REGULATORY CONSIDERATIONS

Incentive-driven FL assigns economic value to data participation, model updates and privacy exposure. Participants should understand how rewards are computed, which metadata are recorded and how disputes are handled. Reputation should not become a permanent penalty without appeal, especially when poor performance is caused by network failure or distribution shift. In regulated domains, only essential information should be stored on-chain; transparency must not become a privacy leak.

11. CONCLUSION AND FUTURE WORK

This paper reviewed representative incentive-driven and blockchain-enabled FL studies and compared them according to contribution evaluation, participant selection, reward allocation, privacy, robustness, and deployability. The comparative results show that existing approaches remain specialized: Shapley and MPC methods provide a strong fairness and privacy interpretation but are computationally expensive; auctions and contracts manage budgets and hidden information but require independent update validation; reputation improves long-term reliability but creates cold-start and calibration risks; and blockchain strengthens auditability and payment enforcement while adding latency and transaction overhead. No reviewed system simultaneously provides budget-aware selection, privacy-preserving contribution evaluation, malicious-update auditing, cold-start mitigation, and atomic settlement in one lifecycle.

Based on these findings, ARIC-FL integrates reputation-aware reverse-auction selection, adaptive contract menus, privacy-preserving multi-factor contribution scoring, commit-reveal validation, encrypted off-chain exchange, and escrow-based atomic settlement within one incentive lifecycle. The framework is evaluated at the design level; future work will implement a prototype and assess its fairness, utility, robustness, scalability, and deployment overhead under IID, non-IID, and adversarial settings.

12. REFERENCES

[1] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and Agüera y Arcas, B. 2017. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 1273-1282. PMLR.

[2] Yang, Q., Liu, Y., Chen, T., and Tong, Y. 2019. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology, 10(2), Article 12. https://doi.org/10.1145/3298981

[3] Ali, A., Ilahi, I., Qayyum, A., Mohammed, I., Al-Fuqaha, A., and Qadir, J. 2023. A systematic review of federated learning incentive mechanisms and associated security challenges. Computer Science Review, 50, 100593. https://doi.org/10.1016/j.cosrev.2023.100593

[4] Zeng, R., Zeng, C., Wang, X., Li, B., and Chu, X. 2021. A comprehensive survey of incentive mechanism for federated learning. arXiv. https://arxiv.org/abs/2106.15406

[5] Chen, K., Zhang, J., Liu, X., Feng, Z., and Yang, X. 2025. piFL: Private, atomic, incentive mechanism for federated learning based on blockchain. Blockchain: Research and Applications,6,100271. https://doi.org/10.1016/j.bcra.2024.100271

[6] Qu, Y., Uddin, M. P., Gan, C., Xiang, Y., Gao, L., and Yearwood, J. 2022. Blockchain-enabled federated learning: A survey. ACM Computing Surveys, 55(4), Article 70. https://doi.org/10.1145/3524104

[7] Gao, L., Li, L., Chen, Y., Xu, C., and Xu, M. 2022. FGFL: A blockchain-based fair incentive governor for federated learning. Journal of Parallel and Distributed Computing, 163, 283-299. https://doi.org/10.1016/j.jpdc.2022.01.019

[8] Behera, M. R., Upadhyay, S., and Shetty, S. 2022. Federated learning using smart contracts on blockchain, based on reward driven approach. arXiv. https://arxiv.org/abs/2107.10243

[9] Zhang, Z., Dong, D., Ma, Y., Ying, Y., Jiang, D., Chen, K., Shou, L., and Chen, G. 2021. Refiner: A reliable incentive-driven federated learning system powered by blockchain. Proceedings of the VLDB Endowment, 14(12), 2659-2662. https://doi.org/10.14778/3476311.3476313

[10] Ouyang, K., Yu, J., Cao, X., and Liao, Z. 2023. Towards reliable federated learning using blockchain-based reverse auctions and reputation incentives. Symmetry, 15(12), 2179. https://doi.org/10.3390/sym15122179

[11] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. 2021. Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1-2), 1-210. https://doi.org/10.1561/2200000083

[12] Xuan, S., Wang, M., Zhang, J., Wang, W., Man, D., and Yang, W. 2024. An incentive mechanism design for federated learning with multiple task publishers by contract theory approach. Information Sciences,664,120330. https://doi.org/10.1016/j.ins.2024.120330

[13] Han, Y., Muazu, T., Samuel, O., and Miao, S. 2025. A federated learning-based selection and incentive system using blockchain technology. Pervasive and Mobile Computing, 112, 102091. https://doi.org/10.1016/j.pmcj.2025.102091

[14] Nguyen, D. C., Ding, M., Pham, Q. V., Pathirana, P. N., Le, L. B., Seneviratne, A., Li, J., Niyato, D., Dobre, O. A., and Poor, H. V. 2021. Federated learning meets blockchain in edge computing: Opportunities and challenges. IEEE Internet of Things Journal, 8(16), 12806-12825.

[15] Toyoda, K., and Zhang, A. N. 2019. Mechanism design for an incentive-aware blockchain-enabled federated learning platform. In Proceedings of the 2019 IEEE International Conference on Big Data, 395-403. IEEE. https://doi.org/10.1109/BigData47090.2019.9006344

[16] Shayan, M., Fung, C., Yoon, C. J. M., and Beschastnikh, I. 2021. Biscotti: A blockchain system for private and secure federated learning. Transactions on Parallel and Distributed Systems, 32(7), 1513-1525.

[17] Weng, J., Weng, J., Zhang, J., Li, M., Zhang, Y., and Luo, W. 2021. DeepChain: Auditable and privacy-preserving deep learning with blockchain-based incentive. IEEE Transactions on Dependable and Secure Computing, 18(5), 2438-2455.

[18] Kang, J., Xiong, Z., Niyato, D., Xie, S., and Zhang, J. 2019. Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory. IEEE Internet of Things Journal, 6(6), 10700-10714.

[19] Le, T. H. T., Tran, N. H., Tun, Y. K., Nguyen, M. N. H., Pandey, S. R., Han, Z., and Hong, C. S. 2021. An incentive mechanism for federated learning in wireless cellular networks: An auction approach. IEEE Transactions on Wireless Communications,20(8),4874-4887. https://doi.org/10.1109/TWC.2021.3062708

[20] Zhan, Y., Li, P., Qu, Z., Zeng, D., and Guo, S. 2020. A learning-based incentive mechanism for federated learning. IEEE Internet of Things Journal, 7(7), 6360-6368.