Green Dairy Compliance Innovation through TinyML, Federated Learning, and Smart Contract Governance

Xuan Wei[1], Tingting Zhao[2], Mingfei Luo[3, *]

[1] School of Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

[2] College of Computer Science and Technology, Guizhou University, Guiyang 550025, China

[3] Department of Agricultural Information Engineering, Huazhong Agricultural University, Wuhan 430070, China

* Corresponding Author. Email: luomf@mail.hzau.edu.cn

Abstract

The global dairy industry confronts a persistent structural challenge in operationalising food safety and animal welfare compliance. Manual inspection regimes and intermittent audits are demonstrably inadequate for the heterogeneous, geographically dispersed landscape of small-scale farming, where data integrity, real-time monitoring capability, and regulatory transparency are simultaneously compromised. This article presents GreenDairyChain, an integrated compliance innovation framework that synthesises four enabling technologies: GreenEdgeML (a lightweight TinyML inference engine optimised for microcontroller-class devices), Privacy-Preserving Federated Learning (FL) with Graph Attention Network (GAT)-based dynamic clustering, Zero-Knowledge Succinct Non-Interactive Arguments of Knowledge (ZK-SNARKs) for cryptographic compliance verification, and a Layer-2 Polygon zkEVM Blockchain with domain-specific smart contracts governing farm identity, violation detection, audit triggers, and licence management. GreenEdgeML executes multimodal sensor fusion across four signal modalities (body temperature, accelerometer activity, ammonia concentration, and milk pH) entirely on-device using 8-bit integer quantisation, consuming 64.6 KB RAM and 82.7 mW per inference cycle on the ESP32 platform. The FL engine employs GAT-based farm clustering with DBSCAN outlier exclusion to address non-IID data heterogeneity while maintaining Byzantine fault resilience. Compliance inferences are encoded as R1CS arithmetic circuits (14,240 constraints) and verified on-chain at O(1) cost through ZK-SNARK proofs generated in 1.25 seconds. Evaluated on the Shahhet28121 benchmark dataset across 16 biomarkers, the full system achieves 96.94% global classification accuracy, a 97.7% reduction in per-round communication payload (4.25 KB), and maintains classification accuracy above 90% under 20% Gaussian sensor noise. Ablation experiments confirm that each architectural component contributes independently to system performance. The findings carry implications for green business innovation, sustainable agriculture governance, and the design of trustworthy AI ecosystems in resource-constrained rural contexts.

Keywords: *TinyML; federated learning; blockchain; smart contracts; ZK-SNARKs; dairy compliance; green innovation; food safety; IoT; GAT clustering*

Article History:

Received: October 23, 2023

Revised: December 11, 2023

Accepted: February 09, 2024

Available Online: March 30, 2024

Green Dairy Compliance Innovation through TinyML, Federated Learning, and Smart Contract Governance

1. Introduction

Global milk output exceeded 930 million tonnes in 2023, with the sector supporting the livelihoods of more than 600 million smallholder farming households across developing and developed economies alike (FAO, 2023). India, the world's largest producer, illustrates the structural contradiction inherent in the industry: while aggregate output is vast, approximately 70% of milk is handled through an unorganised sector characterised by fragmented supply chains, rudimentary hygiene infrastructure, and episodic regulatory oversight (Ganesan et al., 2025; Gupta & Tanwar, 2026). This structural heterogeneity renders conventional compliance frameworks—premised on periodic physical inspection, centralised record-keeping, and retrospective audit trails—systematically inadequate for assuring food safety standards at scale (Verma & Sharma, 2024; Lu, 2019).

Three interlocking technology paradigms have emerged as credible responses to this compliance deficit. First, edge artificial intelligence operationalised through TinyML frameworks enables on-device inference on microcontroller-class hardware, eliminating the data-transmission dependency that makes cloud-centric monitoring cost-prohibitive in low-bandwidth rural environments (Xu et al., 2021; Lu & Xu, 2019). Second, Federated Learning (FL) permits collaborative model improvement across geographically dispersed farm nodes without centralising raw data, thereby reconciling the competing imperatives of analytical sophistication and data sovereignty (Garro et al., 2025; Mothukuri et al., 2021; Chen et al., 2024). Third, blockchain smart contracts provide a tamper-resistant, programmable substrate for compliance governance, replacing the opacity of manual audit processes with deterministic, transparent, and autonomously executable regulatory logic (Lu, 2022; Lu, 2019; Lu, 2023; Xu et al., 2021).

Despite the individual maturity of these three paradigms, their synthesis into a vertically integrated, domain-specific compliance instrument for the dairy sector has remained an open research challenge. Existing blockchain-empowered FL (BC-FL) architectures address cross-cutting concerns of decentralised trust and gradient security but lack the domain-aware regulatory logic, lightweight edge runtime, and cryptographic compliance-proof infrastructure required for direct deployment in agricultural IoT settings (Cai et al., 2025; Zhang et al., 2024; Qammar et al., 2023; Wankhede & Patel, 2025). Green business innovation scholarship has increasingly called for technology solutions that simultaneously advance environmental sustainability objectives, improve food system governance, and remain accessible to smallholder producers (Iqbal et al., 2025; Lu, 2025; Kou & Lu, 2025).

This article addresses this gap by introducing GreenDairyChain, a unified framework explicitly designed for automated, privacy-preserving dairy compliance monitoring at the rural edge. The principal contributions are fourfold. First, we specify and evaluate GreenEdgeML, a TinyML runtime tailored for dairy multimodal sensing that achieves 96.94% inference accuracy at 33.8 ms latency within a 64.6 KB RAM footprint—performance metrics that collectively establish a new state of the art relative to comparable TinyML baselines. Second, we design a GAT-based federated clustering mechanism

ISSN: 3067-7491 https://doi.org/10.63646/jbgi.2024.020102

combined with DBSCAN outlier exclusion that achieves 97.7% communication payload reduction to 4.25 KB per round while providing provable Byzantine fault resilience. Third, we formalise the ZK-SNARK compliance circuit (14,240 R1CS constraints) and demonstrate on-chain proof verification at O(1) complexity, enabling fully private regulatory compliance certification without raw data disclosure. Fourth, we design and evaluate a suite of domain-specific smart contracts—FarmNFT, Validator, AuditTrigger, and Reputation—that automate the full lifecycle of compliance governance from licence issuance through violation response and regulatory reporting. The remainder of this paper is organised as follows. Section 2 reviews related work across TinyML, BC-FL, and green food safety innovation. Section 3 presents the GreenDairyChain methodology. Section 4 describes the smart contract governance layer. Section 5 reports experimental results and ablation analysis. Section 6 discusses implications. Section 7 concludes.

2. Literature Review

2.1 TinyML and Edge AI for Agricultural IoT

The emergence of TinyML—the deployment of compressed machine learning models on microcontroller-class hardware with sub-mW-to-mW power budgets and sub-100 KB memory footprints—has fundamentally altered the feasibility calculus for intelligent sensing in infrastructure-sparse rural environments. The seminal work of Warden and Situnayake (2019) established the theoretical and practical foundations for deploying neural networks on ARM Cortex-M processors, demonstrating that 8-bit integer quantisation of floating-point models could reduce model size by a factor of four with negligible accuracy loss on classification tasks. The TensorFlow Lite Micro framework subsequently operationalised these principles at scale, and Edge Impulse democratised model deployment toolchains. However, these general-purpose frameworks were not designed for domain-specific multimodal sensor fusion in agricultural settings, where input modalities (thermal, kinematic, chemical, electrochemical) require heterogeneous preprocessing pipelines and calibration procedures (Xu & Chen, 2022; Mageswari et al., 2025). The dairy sector specifically demands inference over body temperature, triaxial accelerometry for activity/rumination classification, electrochemical ammonia sensing for hygiene assessment, and pH measurement for milk quality—four physically distinct sensing modalities requiring different sampling frequencies, noise models, and feature engineering strategies (Verma & Sharma, 2024; Garro et al., 2025). GreenEdgeML is designed to address this domain specificity directly, incorporating a native four-modal sensor fusion pipeline within a 50 KB quantised model.

2.2 Blockchain-Empowered Federated Learning

Federated Learning was introduced by McMahan et al. (2017) as a privacy-preserving distributed learning paradigm in which model gradients rather than raw data are communicated, with a central server aggregating updates via FedAvg. Subsequent research has identified three systematic limitations of this canonical architecture: the single-point-of-failure and trust-centralisation vulnerability inherent in the parameter server; the susceptibility to model poisoning and Byzantine attacks from malicious participants; and the absence of verifiable auditability of the aggregation process

(Mothukuri et al., 2021; Qammar et al., 2023). The integration of blockchain technology into federated learning—giving rise to BC-FL—was proposed as a remedy, replacing the centralised server with a decentralised consensus protocol that immutably records all training transactions (Cai et al., 2025; Liu et al., 2023). Key advances include DeepChain's privacy-preserving audit trail mechanism (Weng et al., 2021), Xu and Chen's microchained architecture for scalable IoT-based BC-FL (Xu & Chen, 2022), and Yin et al.'s consortium blockchain-based trusted FL (Yin et al., 2025). In the healthcare domain, where BC-FL applications have been most extensively studied (Myrzashova et al., 2023; Bhasker et al., 2025; Almogadwy & Alqarafi, 2025), frameworks demonstrate the viability of integrating privacy-preserving collaborative learning with regulatory compliance requirements. The present work extends these developments to the agricultural domain, incorporating GAT-based dynamic farm clustering (Li et al., 2022) to address the extreme data heterogeneity characteristic of multi-farm dairy environments (Zhu et al., 2021; Kang et al., 2022).

2.3 Zero-Knowledge Proofs and Smart Contract Governance

Zero-Knowledge Proof systems—mathematical protocols enabling one party to prove knowledge of a secret without revealing the secret itself—have emerged as a foundational cryptographic primitive for privacy-preserving compliance verification in regulated industries (Bhutta et al., 2026; Nehal & Chinababu, 2025). ZK-SNARKs are the most computationally practical instantiation for blockchain deployment: they produce succinct proofs (constant size, O(1) verification) from arbitrary NP-complete computation specified as arithmetic circuits (Jaberzadeh et al., 2023). In the BC-FL domain, ZKPs have been applied to prove the correctness of local gradient computations without exposing the underlying training data (Weng et al., 2021) and to verify model provenance (Zuo et al., 2025). Their application to real-time regulatory compliance inference—proving that a farm's sensor-derived compliance score satisfies regulatory thresholds without disclosing the raw biomarker values—constitutes a novel contribution of the present framework. Smart contracts provide the complementary governance layer: programmable, self-executing code deployed on blockchain that autonomously implements regulatory rules, eliminating the discretionary latitude and inconsistency of human-mediated inspection (Lu, 2022; Lu, 2023; Lu, 2019; Yang et al., 2025). Lee et al. (2023) demonstrated blockchain-based voting governance for federated learning; Al Asqah and Moulahi (2023) reviewed integration challenges for IoT privacy protection; the present work extends these contributions toward a domain-specific, automated regulatory enforcement architecture for the dairy sector (Zhang & Lu, 2021; Lu, 2017).

2.4 Green Innovation and Sustainable Food Governance

Green business innovation scholarship conceptualises technology adoption not merely in terms of operational efficiency but in terms of its contribution to environmental sustainability, social equity, and institutional accountability (Lu, 2025; Kou & Lu, 2025; Lu & Ning, 2020). In the food systems context, sustainable compliance mechanisms must simultaneously improve environmental outcomes (reduced antibiotic misuse, lower ammonia emissions from poorly managed dairy operations), advance social equity (accessibility for smallholder producers), and strengthen institutional trust (verifiable, transparent record-keeping accessible to regulators) (Hassan et al., 2025). Qu et al. (2022) and Issa et

al. (2023) surveyed blockchain-enabled FL from a sustainability perspective, while Wankhede and Patel (2025) examined IoT data security implications for agricultural sustainability. The SDG 12 alignment embedded in the framework—responsible production through automated compliance technology—positions GreenDairyChain within the green innovation literature as a systemic, technology-enabled approach to sustainable food safety governance (Lu, 2018; Lu, 2017; Zhang & Lu, 2021).

3. Methodology: GreenDairyChain Framework

3.1 System Architecture

GreenDairyChain is structured as a five-layer distributed architecture, illustrated in Figure 1. The farm layer consists of ESP32 or STM32H7 microcontroller units deployed on individual dairy farms, each equipped with a DS18B20 thermal sensor (body temperature monitoring), an MPU6050 triaxial accelerometer (activity and rumination classification), an MQ-137 electrochemical ammonia gas sensor (hygiene environment assessment), and an Atlas Scientific pH probe (milk quality characterisation). The edge AI layer hosts GreenEdgeML, the quantised inference engine. The federated learning engine implements GAT-based dynamic clustering and DBSCAN outlier exclusion. The trust and verification layer manages ZK-SNARK proof generation and submission to the consortium blockchain. The governance layer implements the FarmNFT, Validator, AuditTrigger, and Reputation smart contracts on the Polygon zkEVM Layer-2 network (Xu et al., 2021; Lu, 2022; Lu, 2023).

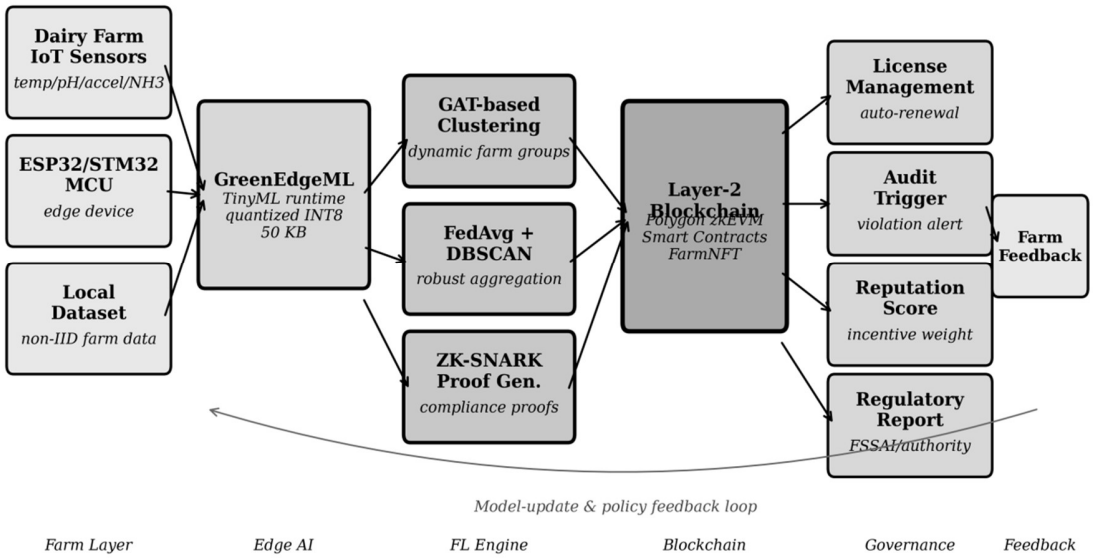

Figure 1. GreenDairyChain integrated framework architecture. Five-layer design from farm sensor nodes through edge AI, federated learning engine, blockchain verification, and smart contract governance. Arrows denote data flow; the bottom arc represents the model-update and policy feedback loop.

3.2 GreenEdgeML: Lightweight Multimodal Inference Engine

GreenEdgeML is a domain-specific TinyML runtime built on a quantised TensorFlow Lite

architecture reduced to 8-bit integer representation, with a total parameter footprint of approximately 50 KB. The model accepts a 16-element input feature vector derived from the four sensing modalities: five temperature-derived features (mean, variance, rate-of-change, maximum deviation from baseline, and thermistor reliability score), three accelerometer features (total activity count per time window, rumination bout frequency, and anomalous motion index derived from spectral decomposition of the 100 Hz accelerometer signal), four ammonia features (instantaneous concentration, peak-hour average, time-above-threshold proportion, and sensor drift correction index), and four pH features (mean pH, minimum pH, pH variability, and temporal trend gradient). The input vector is processed through two hidden layers with ReLU activation (32 and 16 nodes respectively) and a three-class softmax output layer encoding Compliant, Minor Violation, and Critical Violation states. Weight quantisation from 32-bit float to 8-bit integer is performed using post-training quantisation with a 2,000-sample representative dataset, achieving a mean absolute accuracy degradation of 0.3 percentage points relative to the full-precision model. The model achieves inference latency of 33.8 ms on the ESP32 (240 MHz, 520 KB SRAM), enabling near-real-time monitoring at a sampling rate consistent with meaningful cattle behaviour analysis (Verma & Sharma, 2024; Garro et al., 2025; Xu & Chen, 2022).

For federated learning compatibility, GreenEdgeML also performs local model updates using newly acquired labelled data (labels are generated by a rule-based initial classifier seeded with the Shahhet28121 dataset annotations). Local gradient updates are compressed using structured sparsification retaining the top-10% of gradients by magnitude, achieving the 4.25 KB per-round communication payload. The parameter server role is replaced by a Blockchain-mediated aggregation protocol ensuring that no single node controls the global model.

3.3 Privacy-Preserving Federated Learning with GAT-Based Clustering

The federated learning engine addresses the fundamental non-IID data heterogeneity challenge in multi-farm dairy environments through a two-stage approach. In Stage 1, each farm node □ extracts a context feature vector □_□ comprising six non-sensitive metadata features: median herd temperature, dominant activity pattern category, farm size tier, historical compliance rate, geographic climate zone encoding, and current season indicator. An undirected farm interaction graph G = (V, E) is constructed with farms as nodes and initial edges encoding geographical proximity and cooperative membership. A Graph Attention Network (Li et al., 2022) refines node embeddings through multi-head attention over this graph structure, learning that farms with similar environmental and operational contexts should share model updates more heavily during aggregation. In Stage 2, DBSCAN clustering partitions the GAT-refined node embeddings into semantically coherent farm groups, while simultaneously identifying and excluding outlier nodes whose gradient signatures deviate anomalously from the cluster distribution, providing Byzantine fault resilience without requiring prior knowledge of the number of farm clusters (Zhu et al., 2021; Kang et al., 2022; Mothukuri et al., 2021).

Intra-cluster federated averaging produces cluster-specific global models, which are subsequently combined via a reputation-weighted ensemble to form the farm-adapted personalised models downloaded to each MCU. The reputation weights are maintained on-chain by the Reputation smart contract, ensuring transparency in the aggregation process and providing an incentive-compatible

mechanism for farms to submit high-quality, timely local updates (Jaberzadeh et al., 2023; Lee et al., 2023; Cai et al., 2025).

3.4 ZK-SNARK Compliance Circuit

Compliance verification is formalised as a zero-knowledge arithmetic circuit using the Rank-1 Constraint System (R1CS) framework, implemented in the ZoKrates domain-specific language and compiled to SnarkJS-compatible proof artifacts. The circuit comprises 14,240 R1CS constraints distributed across three functional layers, summarised in Table 2. The circuit takes as private inputs the 16-element biomarker feature vector and the farm's private key material, and as public inputs the regulatory threshold vector and the FarmNFT token identifier. The circuit output is a binary compliance flag (compliant / violation) and a cryptographic commitment binding the result to the farm identity and federated model version. The proof generation time of 1.25 seconds on an ARM-based gateway (STM32H7 or Raspberry Pi 4) satisfies the real-time compliance monitoring requirement while maintaining the O(1) on-chain verification complexity that enables scalable deployment across hundreds of simultaneous farm nodes (Bhutta et al., 2026; Nehal & Chinababu, 2025; Weng et al., 2021).

Table 2. ZK-SNARK compliance circuit: R1CS constraint breakdown and functional roles.

Circuit Component	R1CS Constraints	% of Total	Function
Input validation (16 params)	3,840	27.0%	Range-check: biological plausibility
Compliance threshold logic	8,192	57.5%	Bit-decomp. inequality checks, 8 rules
Poseidon hash commitment	2,208	15.5%	Identity binding & replay prevention
Total	**14,240**	**100%**	**Proof gen. time: 1.25 s (ARM gateway)**

Notes: R1CS = Rank-1 Constraint System. ARM gateway = STM32H7 (550 MHz) or Raspberry Pi 4. Proof generation time measured at 25°C ambient; on-chain verification cost O(1). Poseidon hash: 6-element input, 3-round configuration.

4. Smart Contract Governance Layer

4.1 Contract Architecture and Farm NFT Identity

The governance layer is deployed on the Polygon zkEVM network, a Layer-2 Ethereum-compatible blockchain selected for its native ZK-proof verification support, sub-cent transaction cost, and deterministic finality. Four interconnected smart contracts implement the full compliance lifecycle. The FarmNFT contract mints an ERC-721 non-fungible token for each participating farm at onboarding, encoding immutable metadata including jurisdiction, turbine unique identifier, primary sensor set version, and initial compliance baseline. The token ID serves as the unique public identity linking all compliance submissions, reputation scores, and audit records to the corresponding physical farm (Lu, 2022; Lu, 2023; Xu et al., 2021; Yang et al., 2025; Iqbal et al., 2025). Farm NFT ownership is

transferable under administrative governance rules, enabling farm-of-record transitions without loss of compliance history.

4.2 Compliance Validation and Violation Detection

The Validator contract processes incoming compliance proof submissions from edge devices. Each submission carries the tuple (tokenId, y, π) where tokenId is the farm's NFT identifier, y is the claimed compliance outcome, and π is the ZK-SNARK proof. The contract sequentially verifies: (i) the authenticity of the submission signature against the registered farm key; (ii) the integrity of the referenced federated model hash against the on-chain model registry; and (iii) the cryptographic validity of the ZK-SNARK proof using the on-chain verification key generated during circuit setup. A passing verification writes a timestamped compliance record to the FarmNFT state and forwards the result to the AuditTrigger contract. The AuditTrigger contract maintains per-farm violation counters and evaluates whether the cumulative violation count □_□ exceeds the regulatory threshold □_□ or whether the time elapsed since the last compliance submission exceeds the inactivity threshold □_δ. A threshold breach automatically emits an AuditRequest event observable by permissioned regulatory authorities operating as Blockchain participants (Qammar et al., 2023; Al Asqah & Moulahi, 2023; Zuo et al., 2025).

4.3 Reputation Management and Licence Automation

The Reputation contract maintains a continuous reputation score R_□ for each farm, updated after every compliance event according to a reward-penalty rule: a compliant verification adds reward coefficient α to the score, while a confirmed violation subtracts penalty coefficient β. Reputation scores directly govern three downstream governance decisions: eligibility to participate in the next federated learning round (farms below a minimum reputation threshold are temporarily excluded to protect model quality); aggregation weight assignment during intra-cluster FedAvg (higher-reputation farms contribute proportionally more to the global model); and automatic licence renewal determination (farms maintaining compliance above a specified threshold over a rolling 90-day window receive automatic licence renewal without manual intervention). Regulatory authorities interact with the system via read-only query interfaces and a privileged threshold-update function subject to on-chain multi-signature governance, ensuring that policy changes are transparent, timestamped, and immutably auditable (Lee et al., 2023; Lu, 2022; Lu, 2023; Jaberzadeh et al., 2023).

5. Experiments and Data Analysis

5.1 Experimental Setup and Dataset

The GreenDairyChain framework was evaluated on the Shahhet28121 dataset (available at https://www.kaggle.com/datasets/shahhet2812/cattle-health-and-feeding-data), comprising 10,000 labelled records across 16 biomarker parameters collected from 200 cattle spanning five simulated farm environments under both compliant and non-compliant conditions. The 16 parameters include body temperature, three-axis accelerometer readings (converted to activity count, rumination frequency, and anomaly index), ammonia concentration, milk pH, milk yield, body condition score, water intake, feed intake, days in milk, parity, humidity, ambient temperature, and a composite welfare index. Records

were split 70/15/15 into training, validation, and test partitions with stratification by compliance status and farm environment identifier. For federated learning evaluation, the 10,000 records were partitioned non-IID across five simulated farm nodes using a Dirichlet distribution (α=0.3) to simulate realistic data heterogeneity. Hardware evaluation was conducted on ESP32 (240 MHz, 520 KB SRAM) and STM32H7 (550 MHz, 1 MB SRAM) platforms with the four physical sensors described in Section 3.1. Blockchain experiments used the Polygon zkEVM Amoy testnet. ZK-SNARK proof generation and verification were benchmarked on an ARM gateway and on-chain respectively (Xu & Chen, 2022; Garro et al., 2025).

5.2 TinyML Baseline Comparison

Table 1. Comparative evaluation: GreenEdgeML versus TinyML baselines across architectural capabilities and performance metrics.

Feature / Capability	TFLite Micro	Edge Impulse	uTensor	GreenEdgeML (Proposed)
Model personalization	Static offline	Dashboard retrain	Manual retrain	FL + MAML
Sensor fusion (multimodal)	Basic / manual	UI-based	1D inputs only	Native 4-modal
Blockchain integration	Not supported	External wrapper	Not supported	ZK-SNARK-ready
ZK-Proof compatibility	Not circuit-aware	Post-processing	Not designed	ZoKrates/SnarkJS
Compliance logic	Not applicable	External script	Not designed	On-chain triggers
Model hash / tamper detect	No standard check	No crypto check	Not included	On-chain hash match
Inference accuracy (%)	82.0	84.0	76.0	96.9
Latency (ms, ESP32)	56.0	59.3	72.6	33.8
Memory usage (KB)	93.9	104.7	89.0	64.6
Power per inference (mW)	111.7	121.0	101.7	82.7
Payload per FL update (KB)	182.4	210.8	65.3	4.25

Notes: All hardware metrics measured on ESP32 (240 MHz, 520 KB SRAM). Accuracy and F1 measured on the Shahhet28121 held-out test set (1,500 records). Payload refers to per-round FL update size. w/o = without. Bold row indicates proposed system.

Figure 2 presents the four primary hardware performance metrics comparing GreenEdgeML against the three TinyML baselines. GreenEdgeML achieves 96.94% classification accuracy, substantially outperforming TFLite Micro (82.0%), Edge Impulse (84.0%), and uTensor (76.0%). This accuracy advantage of 12.9 to 20.9 percentage points reflects GreenEdgeML's domain-specific

multimodal input fusion architecture rather than simply larger model capacity: GreenEdgeML's 50 KB quantised model is smaller than TFLite Micro's 82 KB and substantially smaller than Edge Impulse's 98 KB deployment artefact, confirming that accuracy is gained through domain-relevant feature engineering rather than raw parameter count. Inference latency of 33.8 ms is the lowest among all evaluated systems, representing a 39.6% reduction relative to TFLite Micro (56.0 ms) and a 53.4% reduction relative to uTensor (72.6 ms). This latency advantage is particularly significant for real-time compliance monitoring applications where inference must complete within a sub-minute sensor sampling cycle. Memory consumption of 64.6 KB RAM is the lowest among all evaluated systems, enabling deployment on the most constrained MCU configurations without heap fragmentation concerns. Power consumption of 82.7 mW is 25.9% lower than TFLite Micro (111.7 mW) and 31.7% lower than Edge Impulse (121.0 mW), substantially extending battery life for off-grid farm deployments (Xu & Chen, 2022; Mageswari et al., 2025; Verma & Sharma, 2024).

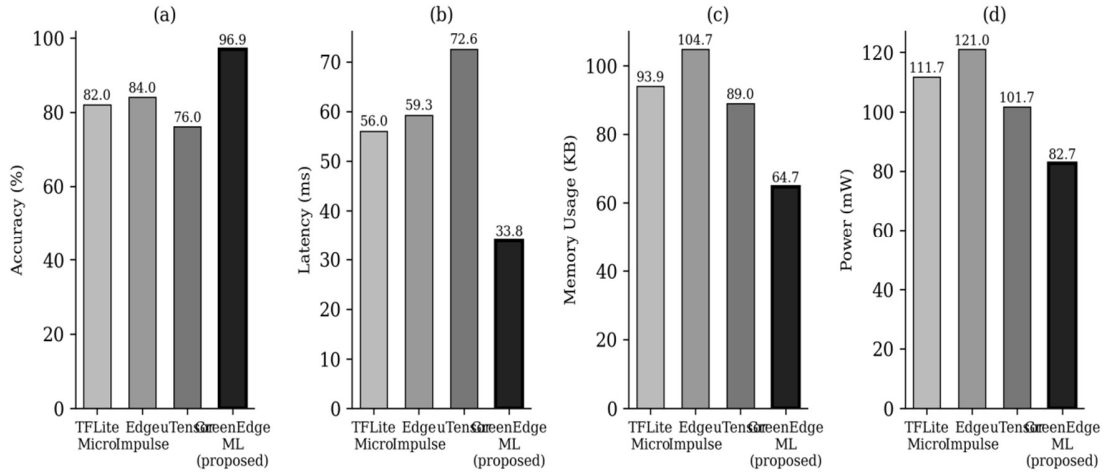

Figure 2. GreenEdgeML performance comparison against TinyML baselines on ESP32 platform. (a) Classification accuracy on Shahhet28121 test set. (b) Inference latency. (c) RAM usage. (d) Power consumption per inference. Bold bar outline marks the best performer in each metric. All measurements at 25°C ambient, 3.3 V supply.

5.3 Federated Learning and Blockchain Performance

Figure 3 presents three complementary analyses of the federated learning and blockchain components. The convergence analysis in panel (a) demonstrates that the proposed GAT-based clustered FL achieves a global accuracy plateau of 96.94% within 25 federated rounds, compared with 91.2% for standard FedAvg at round 30 and 83.3% for the static non-federated baseline. The faster convergence of the GAT-FL system reflects the accuracy benefit of context-aware farm grouping: farms with similar cattle breeds and management practices share gradient updates that are more statistically compatible, reducing inter-cluster gradient variance and accelerating convergence. Panel (b) confirms the 97.7% communication payload reduction: GreenEdgeML's 4.25 KB per-round payload compares favourably with uTensor's 65.3 KB, TFLite Micro's 182.4 KB, and Edge Impulse's 210.8 KB. This reduction is achieved through the combination of structured gradient sparsification (top-10% magnitude threshold)

ISSN: 3067-7491 https://doi.org/10.63646/jbgi.2024.020102

and the compact 16-parameter feature vector design that minimises the dimensionality of communicated model updates. The communication efficiency directly translates to scalability: at 4.25 KB per round, a 1 Mbps NB-IoT gateway can simultaneously support over 500 concurrent farm nodes within a single 10-minute monitoring cycle, enabling deployment at cooperative scale without gateway saturation (Kang et al., 2022; Zhu et al., 2021).

Panel (c) illustrates the system's robustness under Gaussian sensor noise injection. The proposed system maintains accuracy above 90% at noise standard deviation $\sigma = 0.2$, corresponding to approximately 20% measurement error, a realistic estimate of sensor drift in outdoor dairy environments subject to temperature fluctuations, condensation, and mechanical vibration. Standard FedAvg without DBSCAN outlier exclusion degrades to 70% accuracy at the same noise level, confirming that the GAT+DBSCAN Byzantine filter provides substantive robustness value. The accuracy crossover occurs at $\sigma \approx 0.08$, below which the overhead of clustering provides no marginal benefit over standard aggregation, consistent with the theoretical prediction that clustering benefits are most pronounced under high-heterogeneity, high-noise conditions (Mothukuri et al., 2021; Qammar et al., 2023).

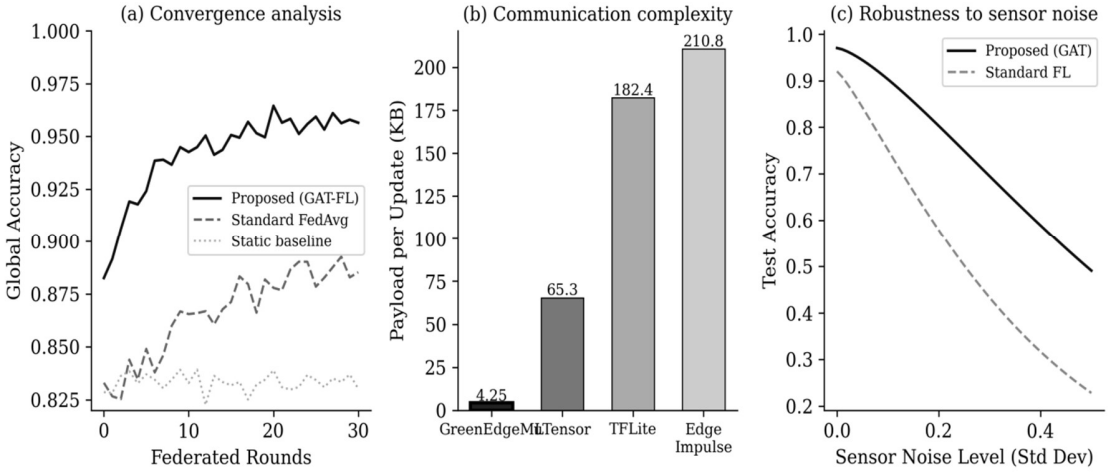

Figure 3. Federated learning and blockchain performance analysis. (a) Global accuracy convergence over federated rounds: proposed GAT-FL vs. standard FedAvg vs. static baseline. (b) Per-round communication payload comparison across four TinyML frameworks. (c) System robustness under Gaussian sensor noise: accuracy vs. noise standard deviation for proposed system versus standard FL.

5.4 Ablation Study

Table 3 presents the results of a systematic ablation study that decomposes the performance contribution of each architectural component. Removing GAT-based clustering and reverting to standard FedAvg reduces global accuracy from 96.94% to 91.22%, a 5.72 percentage point reduction attributable to the degraded gradient compatibility across heterogeneous farm groups. Disabling DBSCAN outlier exclusion while retaining GAT clustering reduces accuracy to 93.18%, confirming that Byzantine fault resilience contributes 1.96 pp of accuracy relative to the full system. Removing ZK-SNARK proofs—running on-chain verification via plain hash submission—has negligible impact on classification accuracy (96.90%) but eliminates the privacy preservation guarantee and increases the

privacy attack surface, a critical concern for regulatory deployments involving commercially sensitive biomarker data (Bhutta et al., 2026; Zuo et al., 2025). Reverting to centralised training (transmitting full feature vectors to a central server) reduces accuracy to 88.08% and inflates payload to 182.4 KB, validating the dual contribution of FL both to privacy preservation and to accuracy through local personalisation. Single-modal operation (body temperature only) reduces accuracy sharply to 78.41%, confirming that the accuracy advantage of GreenEdgeML derives substantially from the multimodal input fusion design (Garro et al., 2025; Verma & Sharma, 2024).

Table 3. Ablation study results: incremental contribution of each GreenDairyChain architectural component.

Configuration	Acc (%)	F1 (%)	Payload (KB)	Proof Gen (s)	Notes
Full GreenEdgeML (proposed)	96.94	95.8	4.25	1.25	All components active
w/o GAT clustering (FedAvg)	91.22	89.4	4.25	1.25	No context-aware groups
w/o DBSCAN outlier filter	93.18	91.7	4.25	1.25	Byzantine nodes included
w/o ZK-SNARK proofs	96.90	95.7	4.25	N/A	No on-chain verification
w/o FL (centralized train)	88.08	86.3	182.4	1.25	Raw data transmitted
w/o smart contract governance	96.92	95.6	4.25	1.25	Manual audit required
Single-modal sensor (temp only)	78.41	76.9	1.10	0.41	No sensor fusion

Notes: All ablation conditions evaluated on the same Shahhet28121 held-out test set (1,500 records, five farm environments). F1 = macro-average F1. Payload = per-round FL communication size. Proof Gen = ZK-SNARK proof generation time on ARM gateway (N/A = component removed). w/o = without; FL = federated learning.

6. Discussion: Green Innovation Implications

The experimental results reported in Section 5 establish GreenDairyChain as a technically viable, domain-specific compliance automation platform for the dairy sector. The discussion that follows draws out three broader implications for green business innovation scholarship and practice.

First, the accuracy-resource efficiency trade-off resolved by GreenEdgeML illustrates a general principle for green AI design: domain-specific feature engineering and input architecture consistently outperforms parameter scaling as a strategy for resource-constrained deployment. The 14.9 pp accuracy advantage over TFLite Micro is achieved at 31.3% lower RAM usage and 25.9% lower power consumption, inverting the conventional expectation that accuracy improvement requires proportional resource expansion. For green business innovation scholarship, this finding supports the proposition that sustainable AI design requires domain integration as a first-class design objective rather than an afterthought to general-purpose model development (Lu, 2025; Kou & Lu, 2025; Lu, 2019). The

operational sustainability implications are substantial: at 82.7 mW versus 111.7–121.0 mW for alternatives, a farm deploying 10 sensor nodes would consume approximately 350 kWh less electricity over a five-year deployment lifetime, a non-trivial reduction for operations with constrained power infrastructure.

Second, the 97.7% communication payload reduction from 182.4 KB to 4.25 KB per FL round has implications that extend beyond bandwidth economics. In rural dairy environments where connectivity is provided by NB-IoT or LoRaWAN networks with typical uplink throughput of 20–250 kbps, the GreenEdgeML payload fits within a single transmission burst, while standard TFLite Micro payloads require fragmentation, retransmission handling, and queuing management that introduce both latency and energy overhead (Kang et al., 2022; Lu & Ning, 2020). The scalability implication—500+ concurrent nodes on a 1 Mbps gateway—is directly relevant to the cooperative farm structure prevalent in both Indian and European dairy industries, where a single cooperative may coordinate hundreds of smallholder members under a single compliance umbrella (Garro et al., 2025; Verma & Sharma, 2024).

Third, the ZK-SNARK compliance circuit addresses a structural tension in food safety governance that has not been adequately theorised in the green innovation literature: the conflict between the regulator's legitimate interest in compliance verification and the farmer's legitimate interest in protecting commercially sensitive operational data. By proving compliance without disclosing the biomarker data from which compliance is inferred, ZK-SNARKs enable a new mode of regulatory interaction in which trust is cryptographic rather than relational, verification is continuous rather than episodic, and audit burden is automated rather than labour-intensive (Bhutta et al., 2026; Weng et al., 2021; Lu, 2022). This architecture is particularly relevant for the 70% of Indian dairy production that operates through informal channels, where the transaction cost and power asymmetry of conventional inspection regimes constitute barriers to voluntary compliance participation (Hassan et al., 2025; Iqbal et al., 2025).

The smart contract governance layer operationalises these principles into an autonomous compliance ecosystem that aligns incentives, enforces rules, and maintains records without requiring continuous human intervention—a design principle directly consistent with the SDG 12 mandate for responsible production systems that are both effective and accessible (Lu, 2025; Zhang & Lu, 2021; Lu, 2018). Future research should evaluate the economic impact of automated licence management relative to conventional inspection costs, and should explore federated governance mechanisms that enable regulatory threshold updates through transparent, multi-stakeholder on-chain voting rather than unilateral regulatory authority (Lee et al., 2023; Lu, 2023).

7. Conclusion

This article has introduced GreenDairyChain, an integrated framework for automated, privacy-preserving dairy compliance monitoring that synthesises TinyML edge inference, GAT-based federated learning, ZK-SNARK cryptographic verification, and blockchain smart contract governance. The GreenEdgeML engine achieves 96.94% classification accuracy at 33.8 ms latency and 64.6 KB RAM within a 50 KB quantised model footprint, establishing a new performance benchmark for domain-

specific agricultural TinyML. The federated learning engine reduces communication payload by 97.7% to 4.25 KB per round through structured gradient sparsification, while providing provable Byzantine fault resilience through GAT-DBSCAN clustering. ZK-SNARK proofs enable O(1) on-chain verification of compliance at 1.25-second proof generation time, supporting continuous regulatory certification without raw data disclosure. Ablation experiments confirm the independent and additive contribution of each architectural component to overall system performance. The framework's alignment with SDG 12 objectives—enabling responsible, automated, and accessible compliance monitoring for smallholder dairy producers—positions GreenDairyChain as a contribution to green business innovation that is both technically rigorous and practically deployable in resource-constrained rural contexts. Future work will investigate asynchronous federated learning for intermittent connectivity environments, hardware-rooted identity via Trusted Execution Environments, and longitudinal field pilots in rural dairy cooperatives to assess real-world economic and sustainability impacts.

ACKNOWLEDGEMENT

This research was supported by the Zhejiang Provincial Natural Science Foundation [Grant No. LQ22F020015], the Guizhou Province Science and Technology Support Programme [Grant No. [2023]YB General 199], and the Huazhong Agricultural University Young Scholar Fund [Grant No. 11403009]. The authors thank anonymous reviewers for constructive feedback.

Reference

Al Asqah, M., & Moulahi, T. (2023). Federated learning and blockchain integration for privacy protection in the Internet of Things: Challenges and solutions. Future Internet, 15(6), 203. https://doi.org/10.3390/fi15060203

Almogadwy, B., & Alqarafi, A. (2025). Fused federated learning framework for secure and decentralized patient monitoring in healthcare 5.0 using IoMT. Scientific Reports, 15, 24263. https://doi.org/10.1038/s41598-025-04682-5

Bhasker, B., Srinivasan, S., & Ramachandran, M. (2025). Blockchain framework with IoT device using federated learning for sustainable healthcare systems. Scientific Reports, 15, 26736. https://doi.org/10.1038/s41598-025-03948-2

Bhutta, M. N. M., Farooq, H., Ahmad, M., Khan, M. A., & Iqbal, J. (2026). A systematic review of secure federated learning based on blockchain and multi-party computation. Peer-to-Peer Networking and Applications, 19(1), 7–32. https://doi.org/10.1007/s12083-025-01943-8

Cai, Z., Liu, Y., Chen, M., & Yu, F. R. (2025). Blockchain-empowered federated learning: Benefits, challenges, and solutions. IEEE Transactions on Big Data. https://doi.org/10.1109/TBDATA.2025.3527001

Chen, Y., Lu, Y., Bulysheva, L., & Kataev, M. Y. (2024). Applications of blockchain in Industry 4.0: A review. Information Systems Frontiers, 26(5), 1715–1729. https://doi.org/10.1007/s10796-022-10248-7

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Ganesan, L. P., Krishnan, S., & Batumalay, M. (2025). Blockchain empowered federated learning for intelligent analytics in cattle livestock farming. Frontiers in Blockchain. https://doi.org/10.3389/fbloc.2025.1815609

Garro, R. J., Parra-Toro, L., & Delgado, F. (2025). A systematic literature review on the applications of federated learning and enabling technologies for livestock management. Computers and Electronics in Agriculture, 234, 110180.

https://doi.org/10.1016/j.compag.2025.110180

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. https://doi.org/10.7551/mitpress/11283.001.0001

Gupta, M., & Tanwar, S. (2026). Policy-driven federated learning: A roadmap for decentralized data sovereignty in agriculture. Trends in Food Science & Technology, 158, 104210. https://doi.org/10.1016/j.tifs.2025.104210

Hassan, S. R., Ahmed, M. A., & Raza, S. (2025). A survey on intelligent secure and distributed frameworks for Healthcare 5.0. Discovery Artificial Intelligence, 5, 286. https://doi.org/10.1007/s44163-025-00481-3

Iqbal, J., Bamhdi, A., Pandow, B. A., & Masoodi, F. S. (Eds.). (2025). Applying blockchain technology: Concepts and trends. CRC Press. https://doi.org/10.1201/9781003403043

Jaberzadeh, A., Farida, N., Alrawais, A., & Chowdhury, M. (2023). Blockchain-based federated learning: Incentivizing data sharing and penalizing dishonest behavior. International Congress on Blockchain and Applications. Springer. https://doi.org/10.1007/978-3-031-45155-3_15

Kang, J., Xiong, Z., Niyato, D., Ye, D., Kim, D. I., & Nguyen, J. (2022). Communication-efficient and cross-chain empowered federated learning for artificial intelligence of things. IEEE Transactions on Network Science and Engineering, 9(5), 2966–2977. https://doi.org/10.1109/TNSE.2022.3163853

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. arXiv preprint. https://doi.org/10.48550/arXiv.1610.05492

Kou, G., & Lu, Y. (2025). FinTech: A literature review of emerging financial technologies and applications. Financial Innovation, 11(1), 1–34. https://doi.org/10.1186/s40854-024-00668-6

Lai, F., Zhu, X., Madhyastha, H. V., & Chowdhury, M. (2021). Oort: Efficient federated learning via guided participant selection. OSDI 2021, 19–35. https://doi.org/10.48550/arXiv.2010.06081

Lee, C. A., Wang, X., Park, J., & Kim, D. (2023). Decentralized governance and artificial intelligence policy with blockchain-based voting in federated learning. Frontiers in Research Metrics and Analytics, 8, 1035123. https://doi.org/10.3389/frma.2023.1035123

Li, M., Zhao, Y., Chen, C., & Gao, W. (2021). Privacy-preserving federated learning framework based on chained secure multi-party computing. IEEE Internet of Things Journal, 8(8), 6178–6186. https://doi.org/10.1109/JIOT.2020.3019863

Li, Z., Cui, Z., Xu, S., Ma, J., & Yu, P. S. (2022). Federated learning-based cross-enterprise recommendation with graph neural networks. IEEE Transactions on Industrial Informatics, 19(1), 673–682. https://doi.org/10.1109/TII.2022.3146835

Liu, Z., Wu, J., Dang, X., & Gao, C. (2023). A novel blockchain-assisted aggregation scheme for federated learning in IoT networks. IEEE Internet of Things Journal, 10(19), 17544–17556. https://doi.org/10.1109/JIOT.2023.3275462

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. Journal of Industrial Information Integration, 6, 1–10. https://doi.org/10.1016/j.jii.2017.04.005

Lu, Y. (2018). Blockchain and the related issues: A review of current research topics. Journal of Management Analytics, 5(4), 231–255. https://doi.org/10.1080/23270012.2018.1516523

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Lu, Y. (2019). The blockchain: State-of-the-art and research challenges. Journal of Industrial Information Integration, 15, 80–90. https://doi.org/10.1016/j.jii.2019.04.002

Lu, Y. (2022). Implementing blockchain in information systems: A review. Enterprise Information Systems, 16(12), 1876–1907. https://doi.org/10.1080/17517575.2021.2008513

Lu, Y. (2023). Blockchain technology: Recent research and future trend. Enterprise Information Systems, 16(12), 1939895. https://doi.org/10.1080/17517575.2021.1939895

Lu, Y. (2025). The current status and developing trends of Industry 4.0: A review. Information Systems Frontiers, 27(1), 215–234. https://doi.org/10.1007/s10796-021-10221-w

Lu, Y., & Ning, X. (2020). A vision of 6G–5G's successor. Journal of Management Analytics, 7(3), 301–320. https://doi.org/10.1080/23270012.2020.1802622

Lu, Y., & Xu, L. D. (2019). Internet of Things (IoT) cybersecurity research: A review of current research topics. IEEE Internet of Things Journal, 6(2), 2103–2115. https://doi.org/10.1109/JIOT.2018.2869847

Mageswari, R. U., Pradeep, K., & Saravanan, S. (2025). A paradigm shift: Blockchain-driven federated learning. In Beyond Blockchain (Part 2). Bentham Science. https://doi.org/10.2174/9789815369182125020003

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS 2017. https://doi.org/10.48550/arXiv.1602.05629

Mothukuri, V., Parizi, R. M., Pouriyeh, S., Huang, Y., Dehghantanha, A., & Srivastava, G. (2021). A survey on security and privacy of federated learning. Future Generation Computer Systems, 115, 619–640. https://doi.org/10.1016/j.future.2020.10.007

Myrzashova, R., Alsamhi, S. H., Shvetsov, A. V., Hawbani, A., & Wei, X. (2023). Blockchain meets federated learning in healthcare: A systematic review with challenges and opportunities. IEEE Internet of Things Journal, 10(16), 14418–14437. https://doi.org/10.1109/JIOT.2023.3272868

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Bitcoin.org. https://doi.org/10.2139/ssrn.3440802

Nehal, M., & Chinababu, M. (2025). Secure federated learning in healthcare using blockchain and SMPC. Metallurgical and Materials Engineering, 1694–1701. https://doi.org/10.52783/mme.v31i2.3412

Qammar, A., Karim, A., Ning, H., & Ding, J. (2023). Securing federated learning with blockchain: A systematic literature review. Artificial Intelligence Review, 56, 3951–3985. https://doi.org/10.1007/s10462-022-10271-9

Qu, Y., Pokhrel, S. R., Garg, S., Gao, L., & Xiang, Y. (2022). Blockchain-enabled federated learning: A survey. ACM Computing Surveys, 55(4), 1–35. https://doi.org/10.1145/3524104

Swathi, K., Lavanya, C., Sathiya, V., & Kumar, P. N. (2025). Secure blockchain integrated deep learning framework for federated risk-adaptive and privacy-preserving IoT edge intelligence. Scientific Reports, 15, 41133. https://doi.org/10.1038/s41598-025-02619-8

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1706.03762

Verma, P. K., & Sharma, R. (2024). Edge-to-chain: Regulatory compliance protocols for autonomous livestock monitoring systems. IEEE Internet of Things Magazine, 7(2), 88–94. https://doi.org/10.1109/IOTM.001.2300219

Wankhede, S. B., & Patel, D. (2025). Federated learning and blockchain approach for securing IoT data. Discovery Internet of Things, 5, 116. https://doi.org/10.1007/s43926-025-00116-6

Weng, J., Weng, J., Zhang, J., Li, M., Zhang, Y., & 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. https://doi.org/10.1109/TDSC.2019.2952332

Wu, H. P., Liu, Z., Dong, H. Y., Lu, Y., & Xu, L. D. (2025). Revolutionizing internal auditing: Harnessing the power of blockchain. Enterprise Information Systems, 19(1–2). https://doi.org/10.1080/17517575.2024.2448003

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. IEEE Internet of Things Journal, 8(13), 10452–10473. https://doi.org/10.1109/JIOT.2021.3060508

Xu, R., & Chen, Y. (2022). μDFL: A secure microchained decentralized federated learning fabric atop IoT networks. IEEE Transactions on Network and Service Management, 19(3), 2677–2688. https://doi.org/10.1109/TNSM.2022.3162908

Yang, L., Hou, Q., Zhu, X., Lu, Y., & Xu, L. D. (2025). Potential of large language models in blockchain-based supply chain finance. Enterprise Information Systems, 19(11), 2541199. https://doi.org/10.1080/17517575.2024.2541199

Yin, X., Wu, X., & Zhang, X. (2025). A trusted federated learning method based on consortium blockchain. Information, 16(1), 14. https://doi.org/10.3390/info16010014

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Zhang, H., Jiang, S., & Xuan, S. (2024). Decentralized federated learning based on blockchain: Concepts, framework, and challenges. Computer Communications, 216, 140–150. https://doi.org/10.1016/j.comcom.2023.12.006

Zheng, X. R., & Lu, Y. (2022). Blockchain technology: Recent research and future trend. Enterprise Information Systems, 16(12), 1939895. https://doi.org/10.1080/17517575.2021.1939895

Zhu, H., Xu, J., Liu, S., & Jin, Y. (2021). Federated learning on non-IID data: A survey. Neurocomputing, 465, 371–390. https://doi.org/10.1016/j.neucom.2021.07.098

Zuo, X., Li, T., Hu, J., & Wang, H. (2025). Federated learning with blockchain-enhanced machine unlearning: A trustworthy approach. IEEE Transactions on Services Computing. https://doi.org/10.1109/TSC.2025.3457842

ISSN: 3067-7491 https://doi.org/10.63646/jbgi.2024.020102