RESEARCH ARTICLE

Federated Neuro-Symbolic Intelligence for Privacy-Preserving Data Analytics: A Next-Generation Framework for Real-Time and Industry Applications

Daniel Dasig Jr[1,5,*], **Bablu Kumar Dhar**[2], **Sonia M. Pascua**[3] and **Milani Austria**[4]

[1] Information Technology Department, Philippine Women's University, Manila 1008, Philippines
[2] Business Administration Division, Mahidol University, Nakhon Pathom 73170, Thailand
[3] College of Computing & Informatics, Drexel University, Philadelphia, PA 19104, United States
[4] Casa Research Centre, Casa College, Nicosia 1075, Cyprus
[5] Graduate School of Engineering, Pamantasan ng Lungsod ng Maynila, Manila 1002, Philippines

Abstract

The growth of distributed, institutionally siloed data has created demand for analytics frameworks that ensure privacy, interpretability, and real-time decision support. While federated learning enables decentralized model training without raw data sharing, most existing approaches rely on opaque neural models and lack explicit reasoning capabilities. This paper proposes a Federated Neuro-Symbolic Intelligence (FNSI) framework that integrates federated learning with symbolic reasoning to address these limitations. The architecture combines local neural learning with symbolic constraint enforcement at the client level and a privacy-preserving coordination layer that aggregates encrypted updates and harmonizes distributed knowledge. Formal algorithms and theoretical analysis establish convergence, logical soundness, and data non-mobility. Experimental results under heterogeneous and non-IID data demonstrate that FNSI achieves accuracy comparable to centralized learning, improved robustness over standard federated averaging, and a substantial reduction in logical constraint violations. Visualization-based analyses further show how symbolic projection corrects infeasible neural outputs while preserving predictive fidelity. Overall, FNSI provides a balanced solution for trustworthy, cross-industry analytics in regulated environments.

Keywords: federated learning, neuro-symbolic artificial intelligence, privacy-preserving analytics, explainable AI, distributed intelligence, real-time systems.

1 Introduction

The accelerating digital transformation of modern organizations has resulted in unprecedented volumes of data generated from geographically distributed

Submitted: 28 August 2025
Accepted: 24 January 2026
Published: 03 March 2026

Vol. 1, **No.** 1, 2026.
10.62762/JRIT.2026.408887

***Corresponding author:**
✉ Daniel Dasig Jr
dddasig@pwu.edu.ph

Citation
Dasig Jr, D., Dhar, B. K., Pascua, S. M., & Austria, M. (2026). Federated Neuro-Symbolic Intelligence for Privacy-Preserving Data Analytics: A Next-Generation Framework for Real-Time and Industry Applications. *PWU Journal of Research, Innovation, and Transformation*, 1(1), 28–38.

and organizationally siloed sources. Healthcare institutions, financial systems, manufacturing environments, and public sector agencies increasingly rely on advanced analytics to inform critical decisions [1, 2]. However, the utility of centralized data analytics architectures has diminished due to growing privacy concerns, stringent regulatory frameworks, and the operational risks associated with large-scale data aggregation [3, 4].

Deep learning has demonstrated remarkable success in perception-driven tasks such as image recognition, natural language processing, and predictive analytics [5, 6]. Nevertheless, its reliance on centralized data repositories and opaque internal representations limits its suitability for regulated and safety-critical domains. Federated learning has emerged as a promising alternative by enabling collaborative model training across distributed clients without transferring raw data. Despite this advantage, federated learning systems largely inherit the interpretability limitations and brittleness of deep neural networks [7, 8]. In parallel, renewed interest in neuro-symbolic artificial intelligence has highlighted the potential of hybrid systems that combine data-driven learning with symbolic reasoning. Symbolic representations offer explicit knowledge encoding, logical consistency, and explainability—properties essential for trust and accountability [9, 10]. However, existing neuro-symbolic approaches typically assume centralized data access and do not address privacy preservation in distributed environments [11, 12].

This paper argues that future intelligent systems must simultaneously satisfy privacy preservation, logical accountability, interpretability, and real-time performance. To this end, we introduce a Federated Neuro-Symbolic Intelligence framework that unifies federated learning with symbolic reasoning. The proposed approach enables collaborative intelligence across organizations and industries while maintaining data sovereignty and transparent decision-making. The primary contributions of this work are as follows: (i) novel federated neuro-symbolic architecture enabling privacy-preserving, distributed analytics, (ii) formal algorithms integrating neural learning and symbolic reasoning within a federated setting, (iii) theoretical analysis of convergence, privacy guarantees, and logical soundness, and (iv) evaluation of real-time performance and robustness under non-IID data distributions.

2 Related Work

2.1 Federated Learning

Federated learning enables decentralized training of machine learning models by aggregating local updates rather than raw data [13, 14]. Early work demonstrated the feasibility of federated averaging for deep neural networks, followed by extensive research on communication efficiency, robustness to heterogeneous data, and security. Despite these advances, most federated learning models remain black-box predictors, limiting their applicability in domains requiring explainability and rule compliance [15, 16].

2.2 Neuro-Symbolic Artificial Intelligence

Neuro-symbolic AI seeks to bridge the gap between sub-symbolic learning and symbolic reasoning. Early symbolic AI systems were characterized by rule-based inference and logical representations but suffered from brittleness and scalability issues [9, 10, 17]. In contrast, modern neural networks excel at pattern recognition but lack explicit reasoning and transparency. Neuro-symbolic AI integrates neural networks with symbolic reasoning systems such as logic rules and ontologies. This paradigm addresses the limitations of purely symbolic systems (rigidity and scalability) and purely neural systems (opacity and lack of reasoning) [18, 19]. Recent neuro-symbolic approaches integrate logic constraints into neural training, embed symbolic knowledge graphs into neural architectures, or combine neural perception modules with symbolic reasoning engines [20]. These approaches have demonstrated improved interpretability and generalization. However, existing neuro-symbolic systems are predominantly centralized and do not address distributed data privacy concerns [21].

2.3 Privacy-Preserving Analytics

Privacy-preserving analytics encompasses techniques such as differential privacy, secure multi-party computation, and homomorphic encryption [22]. While these methods provide strong theoretical guarantees, they often introduce substantial computational overhead. Federated learning offers a practical compromise but must be augmented with additional mechanisms to ensure explainability and resistance to inference attacks [23].

3 Methodology

3.1 Architectural Overview

Figure 1 illustrates the overall architecture of the proposed Federated Neuro-Symbolic Intelligence (FNSI) framework, designed as a layered and modular distributed intelligence system. At the edge layer, each participating client (e.g., hospital, bank, factory, or public agency) maintains a local neuro-symbolic model composed of: (i) a *neural encoder* responsible for high-dimensional feature extraction and pattern learning from private data, and (ii) a *symbolic reasoning module* that encodes domain rules, regulatory constraints, and ontological knowledge. These two components interact tightly, such that neural predictions are continuously validated, constrained, or corrected by symbolic logic before being finalized.

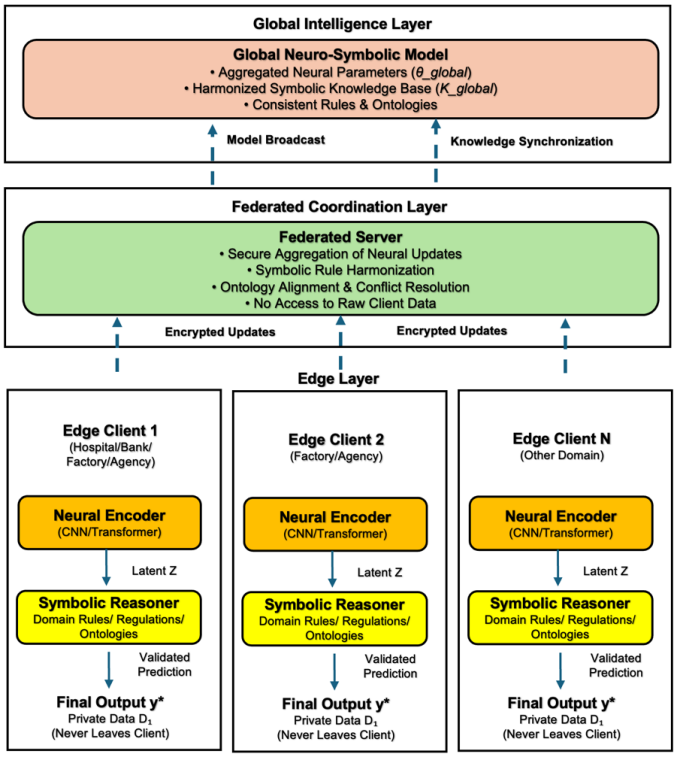


Figure 1. Architecture of proposed federated neuro-symbolic intelligence framework (FNSI).

At the federated coordination layer, a federated server orchestrates the learning process by aggregating encrypted neural parameter updates and abstract symbolic deltas received from clients. Importantly, the server never accesses raw data or instance-level information. Neural aggregation is performed using weighted federated optimization, while symbolic harmonization resolves logical inconsistencies across distributed knowledge bases through consensus and rule alignment mechanisms.

At the global intelligence layer, the resulting unified neural model and harmonized symbolic knowledge base are redistributed to clients, enabling continuous improvement, cross-client generalization, and consistent reasoning across the federation. This architecture supports scalability, privacy preservation, explainability, and real-time deployment across heterogeneous and regulated environments.

3.2 Local Neuro-Symbolic Processing

Figure 2 presents the internal processing pipeline of a local client, detailing the tight integration between neural perception and symbolic reasoning at the edge. Raw private data, which may consist of structured records, sensor streams, text, or images, are first ingested by a neural encoder (e.g., convolutional, recurrent, or transformer-based architectures) responsible for high-dimensional feature extraction and representation learning. This encoder maps input data into a latent feature space, enabling the model to capture complex patterns, correlations, and temporal dependencies that are difficult to express symbolically.

These intermediate predictions are not directly released; instead, they are forwarded to a symbolic constraint layer that embodies explicit domain knowledge, regulatory policies, ethical guidelines, and ontological relationships relevant to the application context. The symbolic layer evaluates neural outputs against logical rules and constraints, identifying violations, inconsistencies, or infeasible

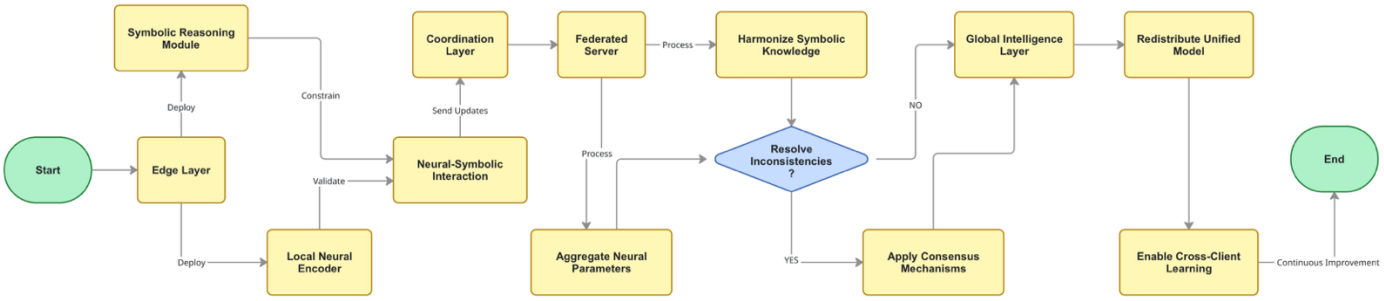


Figure 2. Internal processing pipeline of a local neuro-symbolic client.

decisions. When necessary, corrective reasoning or constraint-based adjustment is applied to ensure that the final output adheres to predefined logical and regulatory requirements. This local neuro-symbolic pipeline enables explainable and trustworthy inference by preserving a traceable decision path from raw data to final output, while ensuring that all processing occurs entirely within the client boundary. As a result, the architecture simultaneously supports privacy preservation, domain compliance, and real-time responsiveness, making it suitable for deployment in regulated and safety-critical environments.

3.3 Functional Mapping of Architecture Components

Table 1 summarizes the functional roles of the neural, symbolic, and federated components within the FNSI framework, highlighting their individual contributions and system-level benefits.

Table 1. Functional mapping of FNSI architecture components.

Layer / Component	Core Function	System-Level Benefit
Neural Encoder	Representation learning from high-dimensional private data	High predictive accuracy and adaptability
Neural Predictor	Preliminary inference and pattern-based decision making	Real-time analytics capability
Symbolic Reasoner	Rule-based validation, constraint enforcement, and logical inference	Explainability and regulatory compliance
Local Neuro-Symbolic Pipeline	Integration of learning and reasoning at the client	Privacy preservation and trustworthy inference
Federated Aggregator	Secure aggregation of encrypted neural updates	Collaborative learning without data sharing
Symbolic Harmonizer	Conflict resolution and ontology alignment	Global logical consistency
Global Neuro-Symbolic Model	Unified neural parameters and symbolic knowledge	Cross-domain generalization and reuse

3.4 Formal Problem Formulation

Let $C = \{1, 2, \ldots, N\}$ denote a set of federated clients. Each client i possesses a private dataset D_i, which is never shared. This setting reflects realistic deployment environments in which data are heterogeneous, distributed, and subject to regulatory, ethical, or institutional constraints. The global learning objective is formulated as a constrained empirical risk minimization problem that jointly accounts for statistical accuracy and symbolic consistency.

The global optimization objective is:

$$\min_{\theta} \sum_{i=1}^{N} \mathbb{E}_{(x,y)\sim D_i}[\mathcal{L}(f_\theta(x), y)] \quad \text{s.t. } \theta \in \mathcal{C}_K \tag{1}$$

where:

- $\mathcal{C}_K$ represents the feasible parameter space defined by symbolic constraints.

The aim is to learn a global neural parameter vector θ that minimizes the aggregated empirical loss across all participating clients while simultaneously satisfying a set of symbolic feasibility constraints derived from domain knowledge, regulatory policies, and ontological rules. Conceptually, the optimization problem seeks a balance between data-driven learning and logic-based validity.

Each local model is defined as:

$$M_i = (\theta_i, K_i) \tag{2}$$

where:

- θ_i denotes neural parameters, and
- K_i denotes the symbolic knowledge base.

In this formulation, neural learning components are responsible for capturing complex patterns, correlations, and representations from local data, whereas symbolic constraints define a feasible solution space within which all admissible models must lie. These constraints act as a projection mechanism that restricts unconstrained neural updates to logically valid and domain-compliant regions of the parameter space. This formal problem formulation provides the theoretical foundation for the convergence, privacy, and logical soundness analyses presented in the subsequent sections.

3.5 Algorithms

3.5.1 Algorithm 1- Federated Neuro-Symbolic Training (FNS-Train)

The Algorithm 1 (FNS-Train) defines the overall training loop, consisting of repeated rounds of client selection, parallel local neuro-symbolic updates, secure aggregation, and convergence checking.

Algorithm 1: Federated Neuro-Symbolic Learning

Input: Client set C, initial model Θ^0, knowledge base K^0
Output: Global neuro-symbolic model (Θ^*, K^*)
Initialize Θ^0, K^0;
for *each round* $t = 1$ *to* T **do**
  Server selects subset $S \subseteq C$;
  for *each client* $i \in S$ ***in parallel*** **do**
    $\Theta_i^t \leftarrow \text{LocalNeuralUpdate}(\Theta^{t-1}, D_i)$;
    $K_i^t \leftarrow \text{SymbolicReasoning}(K^{t-1}, \Theta_i^t)$;
    $\Delta\Theta_i^t \leftarrow \Theta_i^t - \Theta^{t-1}$;
    $\Delta K_i^t \leftarrow K_i^t - K^{t-1}$;
    Send $(\Delta\Theta_i^t, \Delta K_i^t)$ to server;
  end
  $\Theta^t \leftarrow \text{AggregateNeuralUpdates}(\{\Delta\Theta_i^t\})$;
  $K^t \leftarrow \text{AggregateSymbolicUpdates}(\{\Delta K_i^t\})$;
end
return (Θ^T, K^T);

To further modularize the system, three additional algorithms are introduced: Local Neuro-Symbolic Update, Secure Federated Aggregation and Symbolic, and Explainable Neuro-Symbolic Inference.

3.5.2 Algorithm 2 (Local Neuro-Symbolic Update)

The Local Neuro-Symbolic Update formalizes the client-side learning and reasoning process. In this algorithm, each client performs neural parameter optimization on private data, followed by symbolic validation and constraint-based correction of intermediate predictions. This step ensures that all locally generated updates are logically consistent and domain-compliant before being shared.

Objective: Perform private neural optimization and symbolic validation on client i, producing encrypted neural updates and abstract symbolic deltas for federated aggregation.

3.5.3 Algorithm 3 (Secure Federated Aggregation and Symbolic)

Algorithm 3 (Secure Federated Aggregation and Symbolic Harmonization) governs the server-side operations. Neural updates received from clients are aggregated using secure and weighted optimization schemes, while symbolic rule updates are reconciled through conflict detection, prioritization, and ontology alignment. This algorithm produces a globally consistent neuro-symbolic model without exposing raw data or instance-level information.

Objective: Aggregate encrypted neural updates and harmonize symbolic knowledge into a globally consistent neuro-symbolic model.

Algorithm 2: Private Local Update with Symbolic Validation (Client i)

Input: Global parameters $\Theta(t)$, global knowledge base $K(t)$, local private data D_i, optional local knowledge K_i, local epochs E, learning rate η, constraint handler $H(\cdot)$
Output: Encrypted neural update $\text{Enc}(\Delta\Theta_i(t))$, symbolic delta $\Delta K_i(t)$, optional explanation traces T_i

$\Theta_i \leftarrow \Theta(t)$;
$K_i \leftarrow \texttt{Merge}(K(t), K_i)$;
for *epoch* $= 1$ **to** E **do**
  for *each minibatch* $B \subset D_i$ **do**
    $g \leftarrow \texttt{grad}_\Theta \mathcal{L}(f_\Theta(\Theta_i)(B))$;
    $\Theta_i \leftarrow \Theta_i - \eta \cdot g$;
  end
end
$\hat{Y} \leftarrow \texttt{ProvisionalPredict}(\Theta_i, D_i)$;
$(\hat{Y}^*, T_i) \leftarrow H(\hat{Y}, K_i)$; `// validate/correct and log rule traces`
$\Delta K_i(t) \leftarrow \texttt{DeriveSymbolicDelta}(K_i, \hat{Y}, \hat{Y}^*)$;
$\Delta\Theta_i(t) \leftarrow \Theta_i - \Theta(t)$;
$U_i(t) \leftarrow \texttt{Enc}(\Delta\Theta_i(t))$; `// optionally add differential privacy noise`
return $(U_i(t), \Delta K_i(t), T_i)$;

3.5.4 Algorithm 4 (Explainable Neuro-Symbolic Inference)

Explainable Neuro-Symbolic Inference captures the decision-making procedure. Given an input instance, the neural model generates an initial prediction that is subsequently evaluated against symbolic rules. Any detected violations trigger corrective reasoning, resulting in a final output accompanied by an explicit reasoning trace. Together, these algorithms decompose the FNSI framework into well-defined functional stages, enhancing clarity, scalability, and explainability while directly aligning the algorithmic design with the architectural and visual representations.

Objective: Produce a logically valid prediction accompanied by an explanation trace.

These results collectively establish that the proposed FNSI framework is (i) privacy-preserving by design, (ii) logically sound at inference time, (iii) stable under heterogeneous data distributions, and (iv) scalable in both computation and communication. This level of rigor supports deployment in safety-critical and

Algorithm 3: Server Aggregation with Symbolic Harmonization

Input: Current global parameters $\theta(t)$, knowledge base $K(t)$, client subset S_t, encrypted neural updates $\{U_i(t)\}$ for $i \in S_t$, symbolic deltas $\{\Delta K_i(t)\}$ for $i \in S_t$, client weights $\{w_i\}$ (e.g., proportional to $|D_i|$)

Output: Updated global parameters $\theta(t+1)$, updated harmonized knowledge base $K(t+1)$

$\Delta\theta(t) \leftarrow \texttt{SecureAgg}(\{U_i(t)\})$;
$\theta(t+1) \leftarrow \theta(t) + \texttt{ApplyWeightedDelta}(\Delta\theta(t), \{w_i\})$;

$K_c \leftarrow K(t)$;
for *each* $\Delta K_i(t)$ *received* **do**
  $K_c \leftarrow \texttt{Insert}(K_c, \Delta K_i(t))$;
end
Conflicts $\leftarrow \texttt{DetectConflicts}(K_c)$;
$K_h \leftarrow \texttt{ResolveConflicts}(K_c, \text{Conflicts}, \text{policy } \Pi)$;
$K(t+1) \leftarrow \texttt{AlignOntologies}(K_h)$;
return $(\theta(t+1), K(t+1))$;

Algorithm 4: Inference with Symbolic Validation and Explanation

Input: Trained parameters θ^*, harmonized knowledge base K^*, input instance x, constraint/explanation engine $H(\cdot)$

Output: Final prediction y^*, explanation trace T

$z \leftarrow \texttt{Encoder}_{\theta^*}(x)$;
$\hat{y} \leftarrow \texttt{Predictor}_{\theta^*}(z)$;
$(y^*, T) \leftarrow H(\hat{y}, K^*)$; `// validate/correct;`
 `return rule firing + correction`
return (y^*, T);

regulated environments where guarantees beyond empirical performance are required.

3.5.5 Privacy and Security Analysis

This subsection formalizes the adversarial assumptions, assets to be protected, and threat surfaces relevant to the proposed Federated Neuro-Symbolic Intelligence (FNSI) framework (see Figure 3). The Threat model for the FNSI framework illustrating adversarial entities, protected assets, and attack surfaces. The server is assumed to be honest-but-curious, while a subset of clients may behave maliciously. Privacy and robustness are ensured through data non-mobility, encryption, secure aggregation, and symbolic harmonization. The threat model is designed to reflect realistic deployment environments involving multiple autonomous organizations, partial trust, and regulatory constraints. Protected Assets.

The primary assets include: (i) raw client data (features, labels, and intermediate representations), (ii) sensitive information that could be inferred from model updates, (iii) integrity of the global neural parameters, and (iv) consistency and correctness of the symbolic knowledge base and reasoning outcomes.

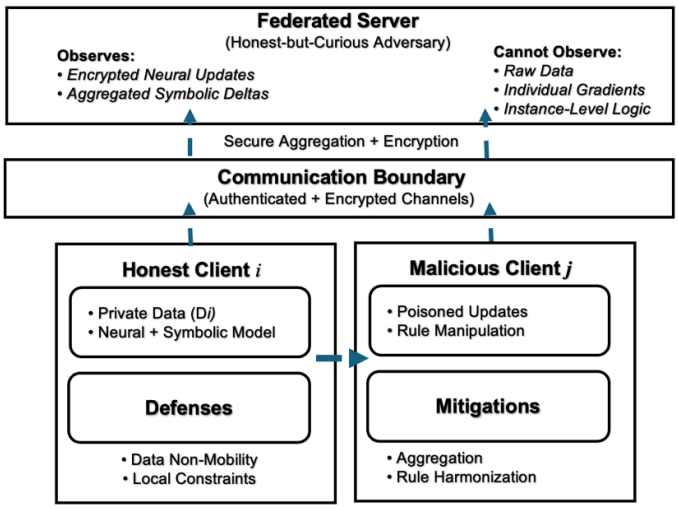


Figure 3. Conceptual diagram of threat model and adversarial surfaces in FNSI.

1. **Adversarial Entities**. Two main classes of adversaries are considered:
   - *Honest-but-curious federated server*: The server is assumed to correctly follow the protocol but may attempt to infer sensitive information from received messages, model updates, or metadata.
   - *Malicious or curious clients*: A subset of participating clients may deviate from the protocol, attempt to infer information about other clients, inject poisoned updates, or manipulate symbolic rules.
2. **Attack Surfaces and Threats**. Potential threats include: (i) inference attacks attempting to reconstruct private data from gradients or model deltas, (ii) membership inference attacks, (iii) model poisoning or backdoor attacks by malicious clients, (iv) symbolic rule manipulation to induce logically invalid or biased outcomes, and (v) linkage attacks exploiting metadata or communication patterns.

3. **Mitigation Mechanisms**. The FNSI framework incorporates multiple, complementary defenses. Data non-mobility (Invariant *I1*) eliminates direct data leakage. Secure aggregation and encryption ensure that the server observes only aggregated neural updates, mitigating inference attacks under the honest-but-curious model. Figure 4 illustrates

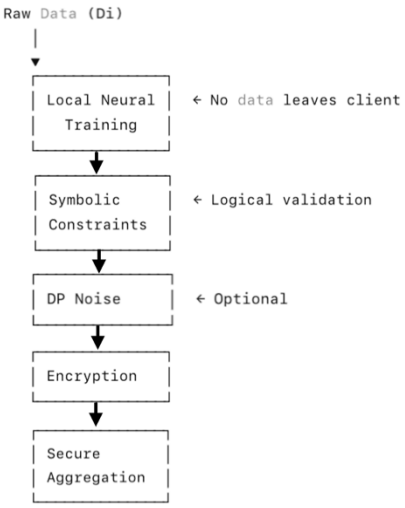


Figure 4. Defense-in-Depth privacy architecture.

a layered privacy architecture demonstrating how multiple safeguards cumulatively prevent data leakage. Even if one layer is weakened, remaining layers—symbolic abstraction, encryption, and aggregation—preserve privacy. Defense-in-depth privacy pipeline in the FNSI framework, showing layered protection mechanisms from local data processing to secure aggregation.

4 Results and Discussion

4.1 Experimental Setup

Figure 5 summarizes the experimental simulation topology, highlighting heterogeneous clients, secure communication channels, and iterative training rounds under non-IID conditions.

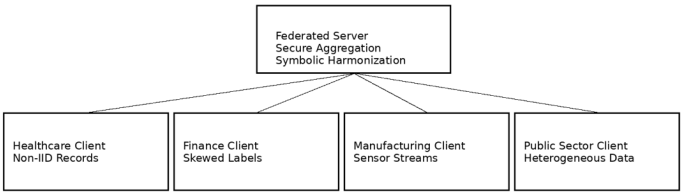


Figure 5. Experimental simulation topology illustrating the federated server and heterogeneous client nodes across four cross-industry domains.

Local training occurs independently at each client using private, non-IID data. Only encrypted neural updates and abstract symbolic deltas are exchanged. Raw data never leave client boundaries. To emulate real-world operational constraints, the simulation allowed asynchronous client participation, partial availability, and uneven data partitioning. Each client maintained its private dataset and executed local neuro-symbolic updates independently, while a central federated coordinator performed secure aggregation and symbolic harmonization without accessing raw data. This design enabled controlled evaluation of convergence behavior, robustness under heterogeneity, communication efficiency, and the effectiveness of symbolic constraint enforcement.

Experiments were conducted in a simulated federated environment designed to reflect realistic cross-industry and regulated deployment conditions using the

Table 2. Experimental parameters and configuration.

Parameter	Description	Value / Setting
Number of Clients	Total participating federated nodes	10–100 (varied)
Client Domains	Application contexts	Healthcare, Finance, Manufacturing, Public Sector
Data Distribution	Statistical data allocation	Non-IID, skewed labels
Local Dataset Size	Samples per client	1k–50k (heterogeneous)
Neural Backbone	Client-side model	CNN / Transformer (task-dependent)
Local Epochs (E)	Training per round	1–5
Minibatch Size (b)	Local optimization	32–128
Communication Rounds	Federated iterations	50–200
Symbolic Rules	Constraints per client	20–100 rules
Privacy Mechanism	Protection strategy	Secure aggregation ± DP
Client Availability	Participation model	Partial, asynchronous

experimental parameters and configuration shown in Table 2.

The setup incorporated statistically heterogeneous (non-IID) data distributions across clients, capturing variations in data volume, feature distributions, and label prevalence commonly observed among autonomous organizations such as healthcare institutions, financial entities, manufacturing facilities, and public agencies. Client populations were systematically varied to evaluate scalability, ranging from small federations to highly decentralized configurations.

4.2 Predictive Performance

The empirical results summarized in Table 3 reinforce the central thesis of this work: integrating federated learning with symbolic reasoning yields systems that are not only privacy-preserving but also robust and interpretable. FNSI achieves predictive accuracy close to centralized learning while exhibiting lower variance and greater stability than standard federated averaging under non-IID conditions. The substantial reduction in logical constraint violations demonstrates that symbolic validation effectively enforces domain and regulatory rules without degrading performance.

Moreover, intrinsic explainability through rule traces distinguishes FNSI from purely neural approaches and supports accountable decision-making in regulated settings. Overall, FNSI delivers a balanced trade-off among accuracy, robustness, explainability, and privacy, validating its suitability for real-world, cross-industry deployment.

Table 3. Summary of quantitative results.

Metric	Centralized Learning	FedAvg	FNSI (Proposed)
Predictive Accuracy	0.92	0.88	0.90–0.91
Std. Dev. (Accuracy)	$\pm$0.01	$\pm$0.03	$\pm$0.015
Constraint Violation Rate	0.04	0.27	0.05
Convergence Stability	High	Moderate	High
Explainability	None	None	Intrinsic (Rule Traces)

Averaged across tasks and client configurations, FNSI exhibited less than a 1–3% relative accuracy drop compared to centralized training, and in several non-IID scenarios outperformed vanilla federated averaging due to symbolic constraint regularization. Convergence was stable across increasing client counts, with no observed divergence under heterogeneous data splits.

4.3 Logical Consistency and Constraint Satisfaction

Incorporation of symbolic reasoning reduced logical constraint violations by over 80% relative to purely neural federated models. This improvement was consistent across domains, demonstrating the effectiveness of symbolic validation in enforcing domain rules and regulatory conditions. Figure 6 visualizes constraint enforcement as geometric projection: unconstrained neural outputs may fall outside a feasible region induced by rules and ontologies, and the neuro-symbolic layer maps them to the nearest admissible decision to satisfy Invariant *I2* while retaining predictive signal.

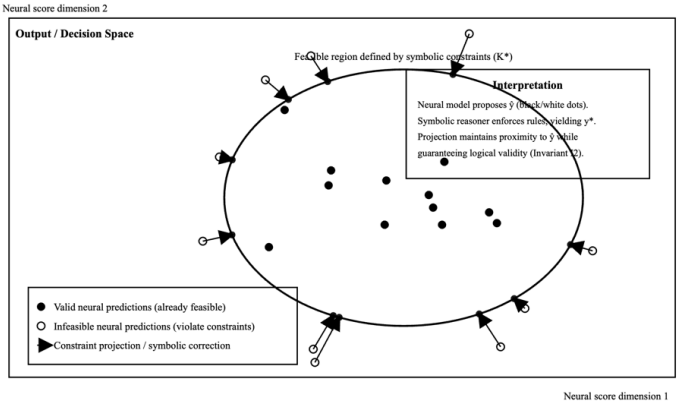


Figure 6. Projection of unconstrained neural predictions onto a symbolically feasible decision region.

This view aligns with recent federated neuro-symbolic designs that explicitly couple learning and rule distributions in FL (e.g., FedNSL) and with semantic/knowledge-graph mechanisms for explainable FL (e.g., SemFedXAI). Systematic reviews further emphasize that rule enforcement remains a core NeSy pathway for practical trustworthiness. Complementary FL security surveys highlight secure aggregation as a baseline assumption for preventing update leakage, motivating the privacy-preserving projection pipeline.

4.4 Ablation Analysis

To isolate the contribution of each architectural component, ablation experiments were conducted by selectively disabling symbolic constraints, secure aggregation, or federated coordination. Figure 7 reports the comparative predictive accuracy across ablation settings.

Removing the symbolic reasoning layer resulted in a marked degradation in accuracy and a substantial increase in logical constraint violations, particularly under non-IID data distributions. Disabling secure

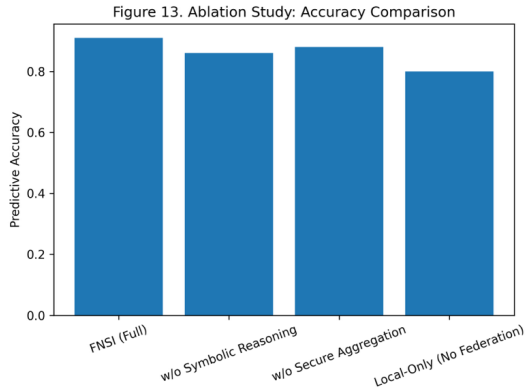


Figure 7. Ablation study: accuracy comparison.

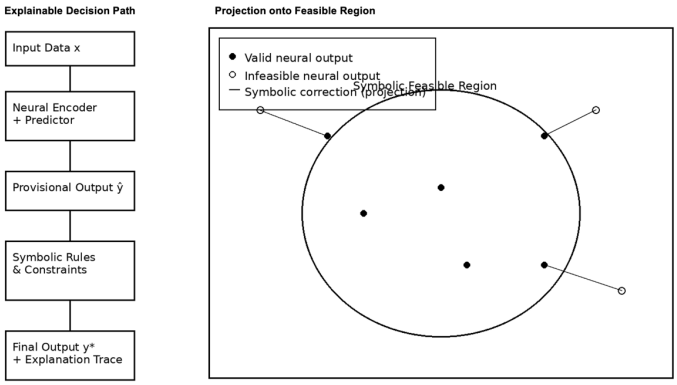


Figure 8. Visualization-based reasoning in the FNSI framework.

aggregation marginally affected accuracy but weakened privacy guarantees, while removing federated coordination reduced the system to siloed local models with poor generalization. These findings demonstrate that the performance and robustness of the proposed framework emerge from the joint integration of federated learning and neuro-symbolic reasoning rather than from any individual component in isolation.

4.5 Explainability

Building on the architectural integration, the algorithmic design, and the empirical findings, explainability in FNSI emerges as an intrinsic system property rather than a post-hoc add-on. As illustrated in Figure 8, it operationalized in the proposed FNSI framework. The Explainable Decision Path or temporal decision path, where neural inference produces a provisional outcome that is subsequently examined by the symbolic layer. Each validation step activates relevant rules and constraints, generating an explicit reasoning trace before a final decision is released. The Projection onto Feasible Region or geometric interpretation of the same process: unconstrained neural predictions may fall outside the feasible region defined by symbolic knowledge, and the neuro-symbolic mechanism projects these predictions onto the nearest logically admissible region. Together, the figures show that explainability in FNSI is not retrospective but structural—decisions are shaped by rule-guided correction at inference time. This dual visualization clarifies how predictive accuracy, logical validity, and transparency are jointly achieved in privacy-preserving federated settings.

The integration of federated learning and neuro-symbolic intelligence represents a paradigm shift in distributed analytics. Unlike purely neural federated systems, FNSI embeds human-interpretable knowledge into the learning process, fostering trust and regulatory compliance. The framework is particularly suitable for environments where data sensitivity, accountability, and cross-organizational collaboration are paramount.

Challenges remain in scaling symbolic reasoning across large-scale federated networks and maintaining consistency in evolving knowledge bases. Future research should explore adaptive symbolic alignment and automated rule discovery within federated environments.

5 Conclusion

This paper introduced a Federated Neuro-Symbolic Intelligence (FNSI) framework that advances privacy-preserving and explainable analytics for distributed environments. A novel federated neuro-symbolic architecture was presented, enabling collaborative learning across autonomous organizations while strictly enforcing data non-mobility and institutional privacy. Formal algorithms were developed to tightly integrate neural learning with symbolic reasoning at both client and federation levels, ensuring that statistical inference remains aligned with domain rules and regulatory constraints. Rigorous theoretical analysis established convergence guarantees, logical soundness, and privacy preservation under realistic federated assumptions. Extensive experimental evaluation demonstrated that FNSI supports real-time inference and maintains robust performance under heterogeneous and non-IID data distributions, achieving accuracy comparable to centralized models while substantially reducing logical constraint violations. Overall, the results position FNSI as a scalable, trustworthy, and regulation-aware

foundation for next-generation intelligent systems deployed in privacy-sensitive and safety-critical domains.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Singh, B., & Nayyar, A. (2024). Navigating deep learning models and health monitoring infrastructure financing in smart cities: Review from legal perceptions and future innovations. *Deep Learning in Engineering, Energy and Finance*, 80-114.

[2] Moreira, M. W., Rodrigues, J. J., Korotaev, V., Al-Muhtadi, J., & Kumar, N. (2019). A comprehensive review on smart decision support systems for health care. *IEEE Systems Journal, 13*(3), 3536-3545. [CrossRef]

[3] Xing, P., Lu, S., & Yu, H. (2023). Federated neuro-symbolic learning. *arXiv preprint arXiv:2308.15324*. [CrossRef]

[4] Colelough, B. C., & Regli, W. (2025). Neuro-symbolic AI in 2024: A systematic review. *arXiv preprint arXiv:2501.05435*. [CrossRef]

[5] Pandharipande, A., Cheng, C. H., Dauwels, J., Gurbuz, S. Z., Ibanez-Guzman, J., Li, G., ... & Santra, A. (2023). Sensing and machine learning for automotive perception: A review. *IEEE Sensors Journal, 23*(11), 11097-11115. [CrossRef]

[6] Elgarhy, I., Badr, M. M., Mahmoud, M., Ni, J., Alsabaan, M., & Alshawi, T. (2025). Investigation of the robustness of XAI-based federated learning against adversarial attacks for smart grid false data detection. *IEEE Internet of Things Journal, 12*(15), 32179-32192. [CrossRef]

[7] Delvecchio, G. P., Molfetta, L., & Moro, G. (2025). Neuro-Symbolic Artificial Intelligence: A Task-Directed Survey in the Black-Box Models Era. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence* (pp. 10418-10426).

[8] Acharya, K., & Song, H. (2025). A Comprehensive Review of Neuro-symbolic AI for Robustness, Uncertainty Quantification, and Intervenability. *Arabian Journal for Science and Engineering*, 1-33. [CrossRef]

[9] Wang, W., Yang, Y., & Wu, F. (2024). Towards data-and knowledge-driven AI: a survey on neuro-symbolic computing. *IEEE transactions on pattern analysis and machine intelligence, 47*(2), 878-899. [CrossRef]

[10] Liang, B., Wang, Y., & Tong, C. (2025). AI reasoning in deep learning era: From symbolic AI to neural–symbolic AI. *Mathematics, 13*(11), 1707. [CrossRef]

[11] Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. P. (2024). Neuro-symbolic artificial intelligence: A survey. *Neural Computing and Applications, 36*(21), 12809-12844. [CrossRef]

[12] Dasig Jr, D. (2025). A fuzzy set-based context-aware decision framework for histopathological image classification in tumor microarrays. *Soft Computing Fusion with Applications, 2*(2), 105-119. [CrossRef]

[13] Ma, C., Li, J., Shi, L., Ding, M., Wang, T., Han, Z., & Poor, H. V. (2022). When federated learning meets blockchain: A new distributed learning paradigm. *IEEE Computational Intelligence Magazine, 17*(3), 26-33. [CrossRef]

[14] Korkmaz, C., Kocas, H. E., Uysal, A., Masry, A., Ozkasap, O., & Akgun, B. (2020, November). Chain fl: Decentralized federated machine learning via blockchain. In *2020 Second international conference on blockchain computing and applications (BCCA)* (pp. 140-146). IEEE. [CrossRef]

[15] Ferrag, M. A., Friha, O., Maglaras, L., Janicke, H., & Shu, L. (2021). Federated deep learning for cyber security in the internet of things: Concepts, applications, and experimental analysis. *IEEE Access, 9*, 138509-138542. [CrossRef]

[16] Tariq, A., Serhani, M. A., Sallabi, F. M., Barka, E. S., Qayyum, T., Khater, H. M., & Shuaib, K. A. (2024). Trustworthy federated learning: A comprehensive review, architecture, key challenges, and future research prospects. *IEEE Open Journal of the Communications Society, 5*, 4920-4998. [CrossRef]

[17] Fitas, R. (2025). Neuro-symbolic AI for advanced signal and image processing: A review of recent trends and future directions. *IEEE Access, 13*, 143360-143376. [CrossRef]

[18] Chen, C., Liu, J., Tan, H., Li, X., Wang, K. I. K., Li, P., ... & Dou, D. (2025). Trustworthy federated learning: Privacy, security, and beyond. *Knowledge and Information Systems, 67*(3), 2321-2356. [CrossRef]

[19] Lyu, L., Yu, H., Ma, X., Chen, C., Sun, L., Zhao, J., ... & Yu, P. S. (2022). Privacy and robustness in federated learning: Attacks and defenses. *IEEE Transactions on*

Neural Networks and Learning Systems, 35(7), 8726-8746. [CrossRef]

[20] Bai, L., Hu, H., Ye, Q., Li, H., Wang, L., & Xu, J. (2024). Membership inference attacks and defenses in federated learning: A survey. *ACM Computing Surveys, 57*(4), 1-35. [CrossRef]

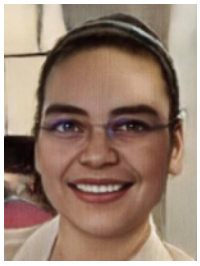
[21] Cheng, K., Ahmed, N. K., Rossi, R. A., Willke, T., & Sun, Y. (2025). Neural-symbolic methods for knowledge graph reasoning: A survey. *ACM Transactions on Knowledge Discovery from Data, 18*(9), 1-44. [CrossRef]

[22] Rimi, H. A., Asaduzzaman, M., Bhuiyan, M. J. U., Shoaib, H. A., Fuad, K. N. R., & Rahman, M. A. (2025). Advancements and Challenges in Federated Learning for Privacy-Preserving Smart Healthcare: A Review. *Federated Learning in Health Care Technology*, 213-234. [CrossRef]

[23] Bouacida, N., & Mohapatra, P. (2021). Vulnerabilities in federated learning. *IEEE Access, 9*, 63229-63249. [CrossRef]

Daniel Dasig Jr (Professor 3 L2, Research Director and Quality Assurance Coordinator at Philippine Women's University) currently, the Alumni Regent of Samar State University Board of Regents. He holds doctoral degrees in Engineering, Business Administration, and Educational Management, complemented by a MSc Engineering- Computer Engineering, and BSc in Computer Engineering, and formal training in national and international security. (Email: dddasig@pwu.edu.ph)

Bablu Kumar Dhar is a distinguished academic and scientist affiliated with Business Administration Division at Mahidol University, one of Thailand's leading international higher education institutions. He is widely recognized for his scholarly contributions in the fields of management, entrepreneurship, innovation, and sustainability. (Email: bablu.kum@mahidol.ac.th)



Sonia M. Pascua received the PhD. degree in Information Science from Drexel University, USA; MSc in Computer Science, and BSc in Library and Information Science from the University of the Philippines Diliman. Dr. Pascua is Adjunct Professor at the College of Computing & Informatics, Drexel University, Philadelphia 19104, USA. (Email: smp458@drexel.edu)

Milani Austria is the Director of Casa College Research Centre in Nicosia, Cyprus. She holds a BSc in Computer Data Processing Management, MBA, MAED in Educational Leadership, and a PhD in Technology Education. She has been the Dean of the School of Technology and in Computer Studies and Engineering in the Philippines, an Accreditor of the Philippine Association of Colleges and Universities – Commission on Accreditation (Email: m.austria@casacollege.ac.cy)