International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Advancing Neurosymbolic AI: A Comprehensive Review of Hybrid Reasoning Frameworks and Applications

Tabish Ali Ansari, Darshan Thube†, Arin Verma‡*

Dept. of Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, Pune, India

*tabish.ansari@aissmsioit.org, †darshanthube8@gmail.com, ‡arinverma2004@gmail.com

ABSTRACT—

Neurosymbolic Artificial Intelligence (NeSy-AI) has emerged as a critical research frontier aiming to integrate the pattern recognition strengths of neural networks with the logical reasoning capabilities of symbolic systems. This review synthe- sizes six recent state-of-the-art contributions that span theoretical foundations, engineering methodologies, security frameworks, industrial robotics, and grounded knowledge representation. Key themes include semantic encoding of symbolic logic into neural networks, domain knowledge-driven anomaly detection, large language model (LLM) augmentation of knowledge graphs, and systematic pattern-based engineering. Despite their varied domains, all works converge toward a common goal: enabling transparent, explainable, and robust AI systems. This paper identifies unresolved challenges, including lack of standardization in system design, evaluation metrics, and real-world deployment readiness, and proposes a conceptual architecture for unifying the field.

***Index Terms*—Neurosymbolic AI, Hybrid AI Systems, Sym- bolic Reasoning, Deep Learning, Knowledge Graphs, Language Models, Explainability, Industrial AI, IoT Security**

1. Introduction

Neurosymbolic AI (NeSy-AI) seeks to bridge the long- standing divide between symbolic reasoning and connectionist learning by uniting rule-based logic systems with deep neural networks. While symbolic AI offers interpretability and com- positionality, neural networks provide adaptability and data-driven learning. This hybrid approach promises to overcome the weaknesses of each paradigm, enabling machines to reason and learn with a level of generality and flexibility previously unattainable.

Recent advancements have demonstrated practical and the- oretical promise, yet the field lacks cohesion in terms of frameworks, design patterns, and evaluation strategies. This review consolidates findings from six diverse but interrelated research articles, offering a panoramic yet coherent view of modern NeSy-AI systems. These contributions are analyzed in terms of their theoretical rigor, engineering sophistication, and domain applicability, with a focus on robotics, cybersecurity, semantic computation, and knowledge representation.

1.A Background and Motivation

Artificial Intelligence (AI) has long been divided into two major paradigms: symbolic AI (or “good old-fashioned AI”) and connectionist AI, most notably neural networks. Symbolic AI, rooted in logic and rule-based systems, emphasizes formal reasoning, structured knowledge representation, and explain-ability. It enables machines to draw inferences from explicitly defined rules, akin to how humans use logic to reason deduc- tively. However, symbolic AI struggles with perceptual data, statistical uncertainty, and scalability to complex real-world environments.

In contrast, neural networks, particularly those used in deep learning, have demonstrated remarkable performance in handling noisy, high-dimensional, and unstructured data—such as images, audio, and natural language. These networks auto- matically learn patterns and latent representations from data, making them invaluable in perception-driven tasks. Yet, they are often criticized as black-box models, lacking transparency, interpretability, and the capacity for abstract reasoning or systematic generalization.

Neurosymbolic AI (NeSy-AI) represents a fusion of these two paradigms—connectionist learning and symbolic reason- ing—to create intelligent systems that not only learn from data but also reason with knowledge. It brings together the learning capability of neural networks with the structured abstraction of symbolic systems, aiming to capture the best of both worlds. The hybrid approach enables machines to handle sensory data and abstract logic simultaneously, mimicking more closely the cognitive processes observed in human reasoning.

1.B What is Neurosymbolic AI?

Neurosymbolic AI combines two computational substrates:

1) **Neural Networks:** These are function-approximators capable of capturing complex nonlinear relationships in data. They are primarily used for pattern recogni- tion, classification, and representation learning. Models like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformers fall within this domain.
2) **Symbolic AI:** These systems are built on explicit rules, ontologies, logic programming, and knowledge graphs. They excel in encoding semantic relationships, perform- ing logical inferences, and offering interpretability.

In NeSy-AI, these components can be combined in several architectural ways:

- Symbolic knowledge embedded into neural network train- ing (e.g., using logical constraints as regularization).
- Neural networks generating symbolic representations (e.g., structured output or logical programs).
- Symbolic systems leveraging neural modules for percep- tion or classification.
- Bidirectional feedback between neural and symbolic components.

For instance, semantic encodings can map symbolic facts into a network's hidden state, while graph-based reasoning mechanisms can be augmented using neural modules for entity disambiguation or relation classification. This synergy enables systems to not only perceive their environment but also to reason about it, explain their decisions, and act with contextual understanding.

1.C Applications of Neurosymbolic AI

The promise of NeSy-AI lies in its broad applicability across data-rich yet knowledge-dependent domains, including:

1) **Knowledge Extraction and Representation:** NeSy-AI enhances automatic extraction of semantic triples from unstructured text and images, which can be stored in knowledge graphs. Systems like Polanyi augment LLM outputs with symbolic graphs for grounded understand- ing.
2) **Explainable AI (XAI):** By aligning neural decisions with symbolic reasoning processes, NeSy-AI systems offer transparent explanations for decisions, critical for domains like healthcare, finance, and security.
3) **Cybersecurity and IoT:** In attack detection, hybrid models leverage neural anomaly detectors and symbolic threat ontologies (e.g., MITRE ATTACK) to classify, contextualize, and respond to threats in real-time.
4) **Robotics and Automation:** In industrial robotics, neu- rosymbolic programming frameworks such as BANSAI help bridge the gap between perception-driven mo- tion planning and symbolic task modeling, facilitating human-readable and adaptable control systems.
5) **Commonsense Reasoning:** NeSy-AI enables the con- struction of Grounded World Models, blending neural commonsense reasoning with ontological structures to understand real-world dynamics and abstract phenom- ena.

These applications underline NeSy-AI's unique position at the confluence of learning, abstraction, and interpretation—a trifecta that most purely neural or symbolic systems fail to simultaneously deliver.

1.D Importance of This Review

The field of NeSy-AI has grown rapidly, yet remains fragmented, with contributions emerging independently across disciplines such as machine learning, semantic web, cyber- security, robotics, and theoretical AI. Many efforts focus on domain-specific implementations without a unifying vision or shared vocabulary. Consequently, it is challenging for researchers to:

- Compare methods rigorously,
- Benchmark systems consistently,
- Design reproducible and interoperable neurosymbolic ar- chitectures.

This review addresses these issues by consolidating and ana- lyzing six foundational and state-of-the-art papers across key application domains. It contributes a systematic synthesis of existing methodologies, a taxonomy of architectural patterns, and a unified conceptual framework. In doing so, it lays the groundwork for future work toward standardization, modular design, and cross-domain applications of NeSy-AI.

1.E Current Progress and Limitations

Recent years have seen important milestones in neurosym- bolic AI:

- Development of formal semantic frameworks for en- coding logic into neural structures (e.g., mapping Horn clauses into fixed-point convergence networks).
- Proposal of pattern-based visual and ontological design tools like SWeMLS, enabling structured engineering of NeSy-AI systems.
- Use of large language models to populate and enrich sym- bolic knowledge graphs dynamically, integrating reactive and generative AI capabilities.

Despite such advancements, critical challenges remain:

1) **Scalability:** Symbolic reasoning can be computation- ally expensive and brittle when scaled to large, noisy datasets.
2) **Interpretability vs. Complexity:** Deep neural models are inherently opaque, and symbolic layers can introduce complexity rather than clarity if not properly integrated.
3) **Lack of Unified Architectures:** Diverse systems lack a consistent architectural grammar, making it difficult to modularize or transfer components across domains.
4) **Benchmarking and Evaluation:** There are few stan- dardized benchmarks for evaluating neurosymbolic sys- tems holistically (in terms of learning, reasoning, and explainability).

2. Literature Review and Analysis

2.A Semantic Foundations of Neurosymbolic Computa- tion

Odense and d'Avila Garcez offer a groundbreaking semantic encoding framework that serves as a theoretical foundation for translating symbolic logic into neural networks. [1] The authors seek to define what it truly means for a symbolic knowledge base to be correctly represented in a neural archi- tecture—a problem historically approached through disparate and often non-generalizable methods. Their framework pro- poses a set-theoretic formalization of the encoding process, wherein each state of a trained neural network corresponds to a logical interpretation. A network qualifies as a neural model of a knowledge base if its state-space converges to the logical models of that base.

This formalism enables encoding Horn clauses, first-order logic, and even probabilistic and fuzzy logic representations into neural networks through architectures that preserve se- mantic integrity. The authors also prove semantic correspon- dence theorems, demonstrating that trained networks can be used not merely for pattern recognition but for semantically coherent logical inference.

The implications of this work are immense: it provides a theoretical backbone for neurosymbolic reasoning systems, ensures logical consistency during inference, and offers a benchmark for evaluating the semantic soundness of hybrid models. Such a framework is indispensable for mission-critical domains where neural decisions must adhere to strict logical rules.

2.B Pattern-Based Engineering of NeSy-AI Systems

Fajar J. Ekaputra presents a methodologically robust and scalable approach to engineering NeSy-AI systems through the development of the SWeML ontology and the concept of pattern-based engineering. [2] By surveying 476 NeSy- AI systems that integrate Semantic Web technologies with Machine Learning (SWeML), the authors develop a knowledge graph of over 40 recurring design patterns, each representing a unique architectural or functional feature.

The framework provides:

- A visual boxology notation for abstract system modeling,
- A layered system description across design, implementa- tion, and operational phases,
- Mechanisms for assessing transparency, auditability, and modularity.

Unlike prior efforts that merely catalog system components, this work offers a formal, ontology-driven standard to docu- ment and compare systems. It parallels the maturity of UML in software engineering, promoting replicability, scalability, and quality assurance in NeSy-AI development. The approach directly supports the interoperability and traceability of hybrid systems and can serve as a foundation for future toolchains in AI system design.

2.C Application in Industrial Robotics: BANSAI

In addressing the AI adoption gap in industrial robotics, Alt et al. propose BANSAI (Bridging the AI Adoption Gap via Neurosymbolic AI), a framework that applies NeSy-AI to real-world robot programming and deployment scenarios. [3] Industrial robotics faces challenges not in perception but in adaptive task planning, human-comprehensible programming, and reusability of domain knowledge. BANSAI introduces a dual representation scheme where:

- Neural components handle low-level sensorimotor func- tions (e.g., grasping),

- Symbolic representations govern task structures and pro- gram synthesis.

BANSAI enables modular robot programming, allowing automation engineers to incrementally integrate AI assistance without disrupting legacy workflows. The framework supports:

- Lifelong learning through surrogate models,
- Program optimization via neural gradients,
- Symbolic reasoning for task planning and safety checks. Most critically, BANSAI allows human-in-the-loop pro- gramming, ensuring interpretability and certifiability. It stands

as one of the first practical systems to deploy neurosymbolic

AI at Technology Readiness Level 6 and beyond, making it a benchmark for applied NeSy-AI in industrial automation.

2.D Security and Explainability in IoT Systems

Kalutharage et al. develop a real-time neurosymbolic intru- sion detection system for Internet of Things (IoT) networks, addressing a domain that demands both speed and explainabil- ity. [4] Their model combines:

- A neural anomaly detection engine (autoencoder-based),
- A symbolic knowledge graph based on the MITRE AT- TACK framework,
- SHAP values for feature attribution and interpretation. This dual-layered model maps low-level statistical anoma-

lies to cyber threat ontologies, enabling the system to infer

not just whether a threat exists, but also its nature, origin, and likely impact on CIA principles (Confidentiality, Integrity, Availability). The system supports:

- Edge-level deployment on devices like Raspberry Pi,
- Explainable decision-making pipelines for cybersecurity analysts,
- Integration with threat intelligence sources for automated response.

Experimental results show over 97% detection accuracy, and its low latency and low compute requirements make it ideal for constrained IoT devices. The framework sets a precedent for actionable and interpretable NeSy-AI in security-critical contexts, bridging AI decision-making with formal, auditable knowledge structures.

2.E Grounded World Models Using LLM-KG Hybrids

De Giorgis et al. introduce Polanyi, a modular neurosym- bolic system that leverages Large Language Models (LLMs) and semantic knowledge graphs (KGs) to create Grounded World Models (GWMs) capable of capturing commonsense and contextual knowledge. [5]

Polanyi operates through a multistage pipeline:

1) Multimodal inputs (text/images) are converted to de- scriptions via LLMs like GPT-4o.
2) Descriptions are transformed into Abstract Meaning Representations (AMR).
3) AMRs are converted into formal OWL2-based knowl- edge graphs using tools like FRED.
4) The resulting KGs are iteratively enriched using heuris- tic rules and feedback from neural systems.

This framework handles implicit knowledge such as:

- Presuppositions,
- Moral judgments,
- Event causality,
- Emotion-based reasoning.

The hybrid architecture falls under Type 2–3 NeSy systems, where symbolic structures act as primary reasoning frame- works, enriched and adapted by neural models.

Applications include:

- Semantic visual reasoning,
- Commonsense inference,
- Natural language understanding,
- Knowledge graph construction from textual or multi- modal inputs.

Polanyi exemplifies how neurosymbolic methods can enable continual knowledge enrichment, a key capability for open- domain AI systems seeking human-like contextual understand- ing.

3. Problem Statement

This paper aims to systematically review and consolidate re- cent advances and propose a modular, explainable framework for hybrid neural-symbolic systems.

4. Research Gaps and Open Challenges

Despite significant progress in the theory and application of neurosymbolic AI, the following core challenges persist across the literature and are hindering widespread adoption and maturity of NeSy-AI systems:

4.A Lack of Standardized Evaluation Metrics

One of the most pressing issues in the NeSy-AI land- scape is the absence of universal benchmarks and evaluation metrics tailored to hybrid systems. While neural network models are traditionally evaluated based on statistical mea- sures like accuracy, F1 score, or mean squared error, and symbolic systems are assessed via logical soundness or infer- ence completeness, neurosymbolic systems operate across both paradigms—requiring new, multifaceted evaluation protocols. For instance, systems like BANSAI or Polanyi output both behavioral predictions (neural) and logical explanations (symbolic), yet no standard exists to evaluate their coherence, fidelity, or alignment between the two reasoning layers. Sim- ilarly, in security-focused systems like Kalutharage et al.'s model, how do we quantify both anomaly detection precision and the semantic validity of the reasoning trace provided by

the knowledge graph?

This creates a barrier to:

- Comparative studies across models,
- Reproducibility of experiments,
- Quantifiable progress in the field.

Until a set of holistic benchmarks is created—possibly in- cluding metrics for interpretability, logical consistency, cross- modal accuracy, and user trust—NeSy-AI will struggle to standardize research outcomes or achieve industry-level con- fidence.

4.B Scalability Bottlenecks in Large-Scale Symbolic Integration

While neural networks scale well due to GPU acceleration and vectorized computation, symbolic components often do not. Logical inference over large ontologies, multi-hop rea- soning in knowledge graphs, or symbolic planning in dynamic environments (like robotics or IoT) can become computation- ally prohibitive.

The BANSAI framework, for example, seeks to embed symbolic reasoning into real-time industrial robot program- ming workflows, which often require millisecond-level reac- tion times and adaptive control over long program sequences. However, integrating full symbolic synthesis or constraint reasoning at such timescales remains technically challenging. Similarly, IoT network security systems must process high- throughput data streams and provide near-instant responses. Introducing symbolic validation layers—such as rule-based MITRE ATTACK reasoning—can introduce latency, rendering the system impractical for deployment on low-power edge devices.

Thus, without advances in:

- Symbolic reasoning parallelization,
- Efficient knowledge representation (e.g., compressed or approximate KGs),
- Hardware acceleration for logic engines,

NeSy-AI systems will be limited to small-scale or offline use cases, curbing their real-world impact.

4.C Limited Reusability and Portability of Components Across Domains

Another significant issue is the lack of modular, reusable components that function reliably across different application areas. Current NeSy-AI systems are often built in silos: frame- works designed for cybersecurity are incompatible with those for robotics, and those created for commonsense reasoning do not generalize to engineering diagnostics.

For example, the symbolic reasoning engine used in Ka- lutharage's cybersecurity model is tailored to map SHAP outputs to MITRE ATTACK ontologies and does not trivially extend to knowledge graphs used in robotics or autonomous vehicles. Similarly, the visual pattern representations intro- duced in the SWeMLS framework offer formal documentation for NeSy-AI architectures but have not yet been adopted as interoperable software blueprints.

This lack of abstraction and decoupling impedes:

- Rapid prototyping of new hybrid systems,
- Cross-domain knowledge transfer,
- Plug-and-play architecture design.

The field urgently needs a repository of domain-agnostic symbolic modules, embeddable neural wrappers, and shared ontologies to facilitate scalable and efficient development.

4.D Absence of Unifying Architectural Blueprints

While individual systems such as Polanyi or BANSAI provide internal architectures, there is no canonical framework that maps out a complete, end-to-end NeSy-AI pipeline—from data ingestion and representation learning to symbolic reason- ing, explanation generation, and feedback adaptation.

The absence of a reference architecture results in:

- Inconsistent terminology and system decomposition,
- Difficulty in onboarding new researchers or practitioners,
- Challenges in evaluating completeness or modularity of implementations.

An ideal blueprint would define:

- Layers (e.g., perception, abstraction, reasoning, explana- tion),
- Interfaces between neural and symbolic modules,
- Reusability hooks and debugging mechanisms,
- Evaluation checkpoints and explainability probes.

Ekaputra's work with pattern-based visual models is a step in this direction but focuses primarily on classification, not unification.

Without such a foundational design guide, neurosymbolic systems risk architectural divergence, which can prevent the formation of a cohesive research and development ecosystem—comparable to what TensorFlow or ROS (Robot Operating System) did for their respective fields.

5. Research Objectives

In light of these gaps, the following research objectives are proposed:

1) To define a modular architectural blueprint for neu- rosymbolic systems combining encoding, reasoning, ex- planation, and retraining loops.
2) To develop domain-agnostic benchmarking protocols for evaluating neurosymbolic performance.
3) To standardize design patterns and visual notations using ontologies like SWeMLS for system documentation and reproducibility.
4) To explore automated composition tools that select neu- ral and symbolic modules based on task specifications.

6. Proposed Layered Neurosymbolic Architec- ture

To unify diverse approaches in neurosymbolic AI, we pro- pose a layered architecture that integrates neural and symbolic components into a coherent, modular framework. Each layer corresponds to a critical function of hybrid intelligence sys- tems:

1) **Data Layer:** *Multimodal input (text, image, network logs).*

This is the input layer where raw data—such as text, images, sensor streams, or network logs—is ingested. It supports multimodal input, reflecting the diverse real- world environments NeSy-AI systems operate in.

2) Neural **Abstraction Layer:** *Embedding and feature extraction using deep networks or LLMs.*

Here, deep learning models or large language models (LLMs) are used to extract features, embeddings, or representations from raw input. This layer performs tasks like object detection, language parsing, or anomaly scoring, transforming unstructured data into usable in- termediate representations.

3) **Symbolic Reasoning Layer:** *Rule-based logic, ontolo- gies, knowledge graphs.*

This layer applies formal reasoning over structured knowledge. It leverages rules, ontologies, and knowledge graphs to perform logic-based inference, plan genera- tion, or verification. It adds interpretability and compo- sitional understanding to the neural outputs.

4) **Explanation Layer:** *SHAP/attention visualization + symbolic alignment (e.g., to MITRE ATTACK or ontol- ogy classes).*

In this layer, tools like SHAP, attention maps, or sym- bolic alignment techniques are used to make the sys- tem's behavior understandable to humans. For example, anomalies can be traced back to violations of symbolic rules, or classifications can be mapped to threat tax- onomies like MITRE ATTACK.

5) **Feedback & Adaptation Layer:** *Looping in symbolic corrections, user feedback, or domain expert input.* This layer enables continual improvement through user or expert feedback. Symbolic inconsistencies can be cor- rected, and the system can be updated with new domain knowledge, enabling lifelong learning and human-in-the- loop alignment.

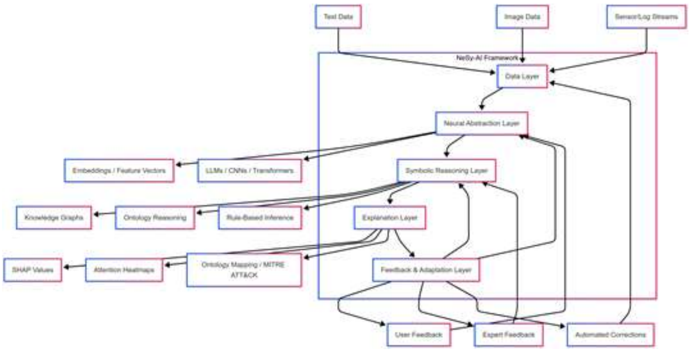

Fig. 1: Layered Neurosymbolic System Architecture with five functional tiers.

Each paper reviewed in this study maps to one or more of these layers:

- Polanyi (De Giorgis et al.) focuses on the Neural Ab- straction, Symbolic Reasoning, and Feedback Layers.
- BANSAI targets Data, Neural, and Symbolic Layers for industrial robotics.
- Kalutharage et al. emphasize Explanation and Symbolic Reasoning for IoT security.
- Ekaputra's pattern-based framework structures and docu- ments systems across all layers.
- Odense & Garcez's semantic framework underpins the Symbolic Reasoning Layer with theoretical rigor.

This architecture provides a blueprint for building modular, scalable, and explainable NeSy-AI systems.

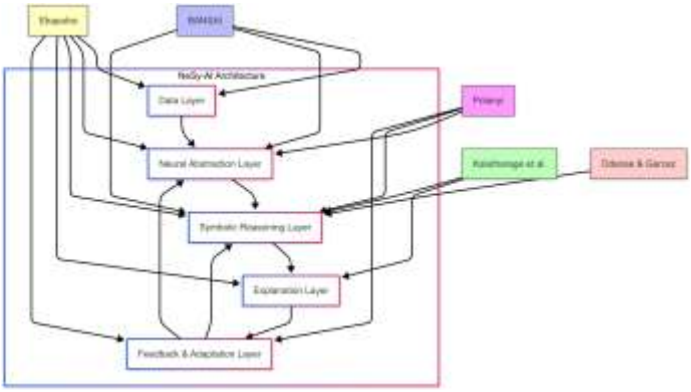

Fig. 2: Layered Neurosymbolic System Architecture with references.

7. Expected Outcomes and Applications

This synthesis and the proposed framework can:

1) Guide future research agendas and system designs in NeSy-AI.
2) Enhance interoperability and reproducibility across neu- rosymbolic applications.
3) Inform real-world adoption in domains such as smart factories, cybersecurity, autonomous systems, and knowledge extraction.

The expected long-term impact includes standards for sys- tem design, more interpretable AI models, and safer AI deployments in high-risk environments.

8. Conclusion

Neurosymbolic AI represents a powerful direction for bridg- ing learning and reasoning in artificial systems. The six papers reviewed provide a microcosm of the field’s diversity and depth—from semantic theory to real-world industrial and security applications. By unifying these contributions into a common architectural and conceptual framework, we hope to accelerate the development of explainable, adaptive, and robust AI systems. Future research must prioritize standardization, scalability, and integration across the symbolic-neural divide.

References

[1] S. Odense and A. d’Avila Garcez, “A semantic framework for neurosym- bolic computation,” *Artificial Intelligence*, vol. 340, 2025.

[2] F. J. Ekaputra, “Pattern-based engineering of neurosymbolic ai systems,” *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 85, p. 100855, 2025.

[3] B. Alt, J. Dvorak, D. Katic, R. Ja¨kel, M. Beetz, and G. Lanza, “Bansai: Towards bridging the ai adoption gap in industrial robotics with neu- rosymbolic programming,” in *Proceedings of the 57th CIRP Conference on Manufacturing Systems (CMS)*, vol. 130, 2024, pp. 532–537.

[4] C. S. Kalutharage, X. Liu, and C. Chrysoulas, “Neurosymbolic learning and domain knowledge-driven explainable ai for enhanced iot network attack detection and response,” *Computers & Security*, vol. 151, p. 104318, 2025.

[5] S. D. Giorgis, A. Gangemi, and A. Russo, “Neurosymbolic graph enrichment for grounded world models,” *Information Processing and Management*, vol. 62, p. 104127, 2025.