Neuro-Symbolic Integration and Knowledge Graphs: The New Frontier of Knowledge Discovery in Theory-Poor Data-Rich Domains (2024–2025)

1. Introduction: The Epistemological Crisis of Data-Rich Science

1.1 The Paradox of "Theory-Poor" Domains

The contemporary scientific and industrial landscape is defined by a fundamental paradox: we possess more data than ever before, yet in many critical domains, our theoretical understanding remains fragmented, provisional, or non-existent. These are the "theory-poor but data-rich" domains. Unlike celestial mechanics or quantum electrodynamics, where robust first-principles mathematical laws (e.g., Schrödinger's equation or General Relativity) constrain the search space of possible models, domains such as systems biology, high-frequency finance, and social network dynamics lack a unifying "Standard Model."

In bioinformatics, for instance, while the Central Dogma of molecular biology provides a high-level framework, the specific regulatory logic governing the interaction of thousands of genes, proteins, and metabolites in a single cell involves stochastic, non-linear feedback loops that defy simple reductionist modeling.[1] Similarly, financial markets are adaptive complex systems driven by millions of interacting agents whose behaviors continuously evolve, rendering static economic laws obsolete almost as soon as they are formulated.[3] In cybersecurity and social network analysis, the "laws" of the system are determined by adversarial strategies and shifting human sociological patterns, creating a non-stationary environment where historical data is often a poor predictor of future states.[5]

1.2 The Failure of Monolithic Paradigms

For the past decade, the dominant response to this data deluge has been Deep Learning (DL). Connectionist architectures, particularly Transformers and Graph Neural Networks (GNNs), have demonstrated an uncanny ability to approximate complex functions and extract features from high-dimensional space. However, as the research from 2024 and 2025 elucidates, pure connectionism faces a "glass ceiling" in theory-poor domains. Neural networks are fundamentally statistical correlators; they approximate the Data Generating Process (DGP) without uncovering the underlying structure. This leads to the "profitability illusion" in algorithmic trading, where models overfit to historical noise [4], or biologically implausible predictions in drug discovery that violate basic chemical constraints.[1] Furthermore, they struggle with "out-of-distribution" (OOD) generalization—failing when encountering entities or scenarios not present in the training set.[7]

Conversely, traditional Symbolic AI—relying on logic, ontologies, and explicit rule systems—offers the interpretability and reasoning capabilities that DL lacks. Yet, purely symbolic approaches are brittle. They require manual knowledge engineering, which is unscalable in the face of petabyte-scale datasets, and they struggle to handle the noise, ambiguity, and continuous variables inherent in real-world observational data.[9]

1.3 The Neuro-Symbolic Convergence (2024–2025)

This report posits that the years 2024 and 2025 mark a decisive inflection point: the maturation of Neuro-Symbolic (NeSy) AI. No longer a theoretical curiosity, NeSy has evolved into a practical necessity for knowledge discovery. The recent literature reveals a shift from "pipelined" hybrids (where a neural net feeds a symbolic solver) to fully unified architectures where symbolic reasoning is differentiable, or where neural perception is constrained by logical axioms during the learning process itself.[10]

Key developments characterizing this period include:

- **The Rise of Graph Foundation Models:** Systems like **GraphOracle** [12] that utilize "Relation-Dependency Graphs" to enable fully inductive reasoning over unseen entities, breaking the "closed-world" assumption of traditional knowledge graph embeddings.
- **Abductive Learning Paradigms:** Frameworks like **ALIGNED** [1] that do not merely classify data but actively generate (abduce) new logical rules to explain empirical anomalies, effectively acting as automated scientists.
- **Quantitative Interpretability:** The formalization of metrics such as **Fidelity** and **Sparsity** [13], moving explainability from a qualitative "nice-to-have" to a rigorous quantitative optimization objective.
- **LLM-Driven Schema Induction:** The use of Large Language Models not as stochastic parrots, but as engines to structure unstructured data into high-fidelity Knowledge Graphs (KGs), which are then used for grounded reasoning.[15]

This report provides an exhaustive analysis of these developments, synthesizing over 130 recent publications to detail how NeSy systems are revolutionizing knowledge discovery in bioinformatics, finance, and social network analysis. We analyze the architectural innovations, the methodological breakthroughs in rule extraction and causal discovery, and the specific applications that validate this new paradigm.

2. Theoretical Foundations: Architectures for Reasoning and Learning

To understand how NeSy systems facilitate discovery, we must first analyze the architectural innovations that allow the fusion of continuous (neural) and discrete (symbolic) representations. The taxonomy of these systems has refined significantly in the 2024–2025 period.

2.1 A Taxonomy of Neuro-Symbolic Integration

Recent surveys [10] propose taxonomies based on the "tightness" of coupling and the functional role of the logic. We can categorize the state-of-the-art approaches into four distinct classes:

2.1.1 Logically Informed Embeddings (Regularization)

In this architecture, logic acts as a soft constraint or a regularizer on the neural loss function. The neural network remains the primary inference engine, but its optimization landscape is shaped by symbolic rules.

- **Mechanism:** Frameworks like **Logic Tensor Networks (LTN)** and **Semantic Regularization** treat logical axioms (e.g., $\forall x: \text{Human}(x) \implies \text{Mortal}(x)$) as differentiable constraints. If the network assigns a high probability to "Human" and a low probability to "Mortal" for the same entity, a penalty is added to the loss.[19]
- **Application:** This is widely used in **Knowledge Graph Completion (KGC)**. By enforcing transitivity or symmetry rules in the embedding space (e.g., vector space constraints in RotatE or TransE), models achieve higher accuracy on link prediction tasks while adhering to domain logic.[9]

2.1.2 Symbolic-Driven Neural Reasoning

Here, the symbolic system directs the neural network's focus. The symbolic component (often a solver like Prolog or an ASP engine) defines the structure of the computation, while neural networks act as "neural predicates" to interpret raw data.

- **Mechanism: DeepProbLog** is the archetype. A logic program defines the reasoning (e.g., mathematical addition), while a neural network recognizes the digits from images. The gradient from the logical result (the sum) backpropagates through the logic program to train the digit recognizer.[19]
- **Advantage:** This allows for learning from weak supervision. The system learns to recognize "2" not because it was labeled "2", but because "2 + 2 = 4" was the only logically consistent interpretation of the equation.

2.1.3 Neural-Driven Symbolic Reasoning (Program Induction)

The neural network operates as a generator, outputting symbolic structures—rules, graphs, or programs—which are then executed by a symbolic engine.

- **Mechanism:** Models utilize sequence-to-sequence architectures to translate natural

language or raw data into **Domain Specific Languages (DSLs)** or logical queries (SPARQL, SQL). In 2025, this has evolved into **Rule Extraction** networks that distill the behavior of a deep network into a set of explicit IF-THEN rules.[20]
- **Significance:** This is crucial for **Interpretability**. The neural network performs the heavy lifting of intuition, but the final decision is made by an auditable symbolic rule, bridging the gap between accuracy and transparency.[7]

2.1.4 Fully Integrated Differentiable Logic

The frontier of 2025 research lies in fully differentiable logic, where logical operations themselves are relaxed into continuous functions (e.g., T-norms in fuzzy logic), allowing for end-to-end training of both the perception and reasoning modules simultaneously.

- **Mechanism: Differentiable Inductive Logic Programming (∂ILP)** allows a system to learn the logical rules *themselves* from data using gradient descent. Unlike traditional ILP, which searches through a discrete space of clauses, ∂ILP learns the weights of logical templates.[12]

2.2 The Emergence of Graph Foundation Models

A critical limitation of earlier NeSy systems was the "Transductive" nature of KG embeddings—they could only reason about entities seen during training. In theory-poor domains like drug discovery or cyber-threat intelligence, the entities (new molecules, new malware variants) are constantly changing. 2025 has seen the rise of **Graph Foundation Models** capable of **Inductive Reasoning**.

2.2.1 GraphOracle and Relation-Dependency Graphs (RDG)

The **GraphOracle** model, introduced in May 2025, exemplifies this shift.[8] It addresses the challenge of generalizing to entirely new graphs with unseen entities.

- **Architecture:** GraphOracle fundamentally shifts the embedding focus from entities to *relations*. It constructs a **Relation-Dependency Graph (RDG)**, which explicitly encodes the compositional patterns and dependencies between relations (e.g., how "lives_in" and "born_in" relate to "citizenship").[12]
- **Mechanism:** By using a query-dependent attention mechanism over the RDG, GraphOracle learns inductive representations of relational topology. When it encounters a new graph, it doesn't need to learn embeddings for the nodes; it recognizes the *structural patterns* of the edges.
- **Performance:** Extensive experiments on 31 benchmarks demonstrate that GraphOracle improves prediction performance by up to **35%** compared to strong baselines like ULTRA or INGRAM in cross-domain settings.[12] This confirms that relational logic is transferable across domains, even when the underlying entities are disjoint.

2.3 The Abductive Learning (ABL) Paradigm

In data-rich, theory-poor domains, the goal is often **Abduction**—inferring the best explanation for the data. Abductive Learning (ABL) formally integrates machine learning with first-order logic to perform this task.[11]

- **The ABL Loop:**
  1. **Perception:** A neural network processes raw data (e.g., images of cells) and predicts pseudo-labels (e.g., "Phenotype A").
  2. **Consistency Check:** A symbolic Knowledge Base (KB) checks if these pseudo-labels are consistent with known background knowledge (e.g., "Phenotype A cannot occur if Gene X is deleted").
  3. **Abduction:** If an inconsistency is found, the system uses logical abduction to find the minimal change to the pseudo-labels or the rules that would restore consistency.
  4. **Optimization:** This "corrected" data is used to retrain the neural network, or the "abduced" rule is added to the KB.
- **Impact:** ABL enables **Neuro-Symbolic Bootstrapping**.[11] It allows the system to learn from unlabeled data by using the logic to self-supervise the neural network. In the **ALIGNED** framework (discussed in Section 4), this approach allows for the *correction* of biological knowledge bases using data, turning the AI into a mechanism for scientific discovery.[1]

2.4 LLMs as Schema Inducers and Reasoners

The integration of Large Language Models (LLMs) into the NeSy stack is a dominant theme of 2025. LLMs are used for **Schema Induction**—automatically defining the ontology (classes and relations) of a Knowledge Graph from unstructured text.[15]

- **Workflow:** An LLM scans a corpus (e.g., medical journals), identifies recurring concepts (Entities), and proposes a schema. A symbolic reasoner then validates this schema for logical coherence (e.g., detecting circular definitions).
- **GraphRAG:** Systems like **KGQAGen** [24] and **GraphRAG** [6] use the constructed KG to ground the LLM's generation. When answering a query, the LLM retrieves structured paths from the KG, ensuring its reasoning is factually supported and reducing hallucination. This combination leverages the LLM's linguistic flexibility with the KG's factual rigidity.

3. Methodological Innovations: From Prediction to Discovery

The transition from predictive AI to discovery-oriented AI requires specific methodological tools. The literature of 2024–2025 highlights three key areas of innovation: Interpretability Metrics, Causal Discovery, and Handling Inconsistency.

3.1 Quantitative Interpretability: Fidelity, Sparsity, and Beyond

In theory-poor domains, a "black box" prediction is often insufficient. If an AI predicts a stock crash or a drug interaction, the *mechanism* is more valuable than the prediction itself. 2025 has seen the formalization of quantitative metrics to evaluate NeSy explanations.[13]

Metric	Definition	Importance in NeSy
Fidelity	The degree to which the symbolic explanation (e.g., rule set) matches the predictions of the underlying neural model.	Ensures the explanation is truthful and not a simplification that hides the model's actual behavior.[13]
Sparsity	The compactness of the explanation, typically measured by the number of rules, predicates, or edges in a subgraph.	A rule set with 1,000 conditions is accurate but not interpretable. Optimization seeks the Pareto frontier between fidelity and sparsity.[14]
Balanced Consistency	A composite metric weighing the agreement of predictions with both empirical data and symbolic background knowledge.	Critical for ABL systems (like ALIGNED) to balance "fitting the data" vs. "adhering to theory".[1]
Rule Complexity	Measures the cognitive load of a rule (e.g., tree depth, number of variables).	Ensures that extracted knowledge is human-consumable.[14]

These metrics are now standard in benchmarking, replacing the subjective "visual inspection" of attention maps that characterized earlier XAI research.[25]

3.2 Causal Discovery in Neuro-Symbolic Systems

Correlation is not causation, a truism that is particularly dangerous in data-rich environments where spurious correlations abound. NeSy systems in 2025 are increasingly deployed for **Causal Discovery**—inferring the directionality of relationships from observational data.[28]

- **Categorical Causal Models:** Frameworks like **Democritus** [29] utilize Category Theory to formalize causal models. They leverage LLMs to propose causal hypotheses (narratives)

from text, which are then rigorously tested against numerical data using statistical independence tests.
- **Intervention Modeling:** Systems like **RealTCD** [30] attempt to reconstruct causal networks from time-series data without explicit interventional experiments (which are impossible in finance or macroeconomics). They rely on the asymmetry of noise distributions and logical constraints to infer causal direction.

3.3 Handling Inconsistency and Noise

Real-world Knowledge Graphs are never perfect. They contain contradictions (e.g., sourced from conflicting papers) and gaps. Traditional symbolic reasoners (like OWL reasoners) often fail catastrophically in the presence of a single logical contradiction.

- **Embedding-Based Reasoners (EBR):** New approaches [31] embed the logical theory into a vector space. In this space, a contradiction is not a fatal error but a region of high "energy" or low probability. This allows the system to reason robustly over inconsistent KGs.
- **Performance:** Experiments show that EBRs maintain high reasoning performance (Jaccard scores) even when 10–20% of the axioms in the KB are corrupted with noise, significantly outperforming crisp symbolic reasoners like Pellet.[32] This robustness is essential for "Knowledge Discovery" where the initial data is inherently messy.

4. Domain Deep Dive: Bioinformatics and Computational Biology

Bioinformatics is the archetypal "theory-poor, data-rich" domain. While we have sequenced genomes (data), the functional logic of the cell (theory) remains largely opaque. NeSy systems are now being used to reverse-engineer this logic.

4.1 Gene Regulatory Network (GRN) Inference via LogicSR

Inferring the control circuits of the cell—Gene Regulatory Networks (GRNs)—is a massive challenge. Standard correlation networks are dense and uninterpretable ("hairballs").

- **Methodology:** The **LogicSR** (Logical Symbolic Regression) framework [2] integrates the mechanistic interpretability of Boolean networks with symbolic regression. It uses a **Multi-Objective Monte Carlo Tree Search (MCTS)** to explore the vast space of possible logical functions (e.g., "Gene A activates Gene B ONLY IF Gene C is absent").
- **Innovation:** By incorporating prior biological knowledge (from databases like GO or KEGG) as a constraint in the MCTS, LogicSR ensures that the discovered networks are biologically plausible.
- **Outcome:** The system discovers sparse, interpretable logical rules that explain cell differentiation pathways (e.g., in glioblastoma) with higher accuracy than pure deep learning methods, which often overfit to noise.[2]

4.2 The ALIGNED Framework and Knowledge Refinement

Genetic perturbation screens (e.g., CRISPR knockouts) provide data on how cells respond to gene deletions. However, existing Knowledge Bases (KBs) are often incomplete.

- **The Framework: ALIGNED** (Adaptive aLignment for Inconsistent Genetic kNowledgE and Data) [1] is an end-to-end ABL framework. It consists of a neural predictor (predicting expression changes) and a symbolic reasoner (encoding the KB).
- **Discovery Mechanism:** ALIGNED introduces a **Balanced Consistency** objective. It doesn't just try to fit the data; it tries to maximize consistency with the KB. However, if the data overwhelmingly contradicts the KB, the system performs **Knowledge Refinement**—it suggests changes to the KB (e.g., adding a missing interaction).
- **Results:** The framework has been shown to "re-discover" known biological mechanisms and, more importantly, propose new regulatory interactions that were validated by literature, effectively automating the scientific hypothesis generation process.[1]

4.3 Drug Repurposing with Explainable Paths

Deep Learning has been used to predict Drug-Disease links, but clinicians require an explanation.

- **System:** Recent work [20] integrates Knowledge Graph Embeddings (KGE) with symbolic rule extraction. The **KGML-xDTD** model uses reinforcement learning to find paths in the KG that explain a predicted link (e.g., Drug X $\to$ Target Protein Y $\to$ Pathway Z $\to$ Disease W).
- **Impact:** This converts the prediction into a testable hypothesis. It also enables **Zero-Shot** predictions for new drugs by reasoning about their structural similarity and interaction profiles, even if no clinical data exists for that specific compound.[35]

5. Domain Deep Dive: Financial Markets and Economic Systems

Financial markets represent a different kind of "theory poverty." The system is adversarial and adaptive; once a "theory" (trading strategy) becomes known, market participants adapt, and the theory ceases to work (the Efficient Market Hypothesis in action). NeSy systems are used here to find transient, evolving logic.

5.1 Trading Rule Discovery via Genetic Network Programming

Pure Deep Reinforcement Learning (DRL) agents often act as black boxes, making them risky for institutional deployment.

- **Methodology:** Researchers have combined DRL with **Genetic Network Programming (GNP)**.[37] Unlike Genetic Algorithms that evolve linear strings, GNP evolves graph-based logic structures.

- **Symbolic Output:** The output of the system is not just a "Buy" signal, but a logical rule: "IF (RSI < 30) AND (Volume > Moving_Avg) AND (Sentiment > Positive) THEN Buy."
- **Advantage:** These rules are interpretable and can be audited by risk managers. They can also be checked against regulatory constraints (e.g., "Do not trade if volatility > X"). The NeSy approach outperforms pure DRL in learning asset-specific rules that are robust to market regime changes.[37]

5.2 Combating the "Profitability Illusion"

A major finding in 2025 is the prevalence of the **"Profitability Illusion"** in AI-driven finance—where models appear profitable in backtests due to overfitting noise or ignoring execution costs.[4]

- **Solution:** Protocols like **AutoQuant** [4] utilize a neuro-symbolic approach to rigorous configuration and auditing. Instead of treating the AI as a "strategy generator," it treats it as a "configuration selector" under strict symbolic constraints (transaction costs, liquidity limits, window semantics).
- **Inference:** The system forces the "fitting assumptions" to be explicit and auditable, preventing the discovery of spurious "money-printing" machines that fail in live trading.

5.3 Financial Knowledge Graphs (FKG) and Systemic Risk

Modeling systemic risk requires understanding the intricate web of interdependencies between firms, supply chains, and obligations.

- **Reasoning Paths:** Systems utilizing **Vadalog** (a Datalog extension) on Financial Knowledge Graphs (FKGs) allow for complex reasoning.[39] For example, determining "Risk of Contagion" involves multi-hop reasoning: "If Firm A defaults, and Firm B has high exposure to A, and Firm C is a supplier to B, what is the risk to C?"
- **Causal Discovery:** Tools like **Democritus** [29] are used to build **Large Causal Models (LCMs)** of the market. By parsing millions of financial documents with LLMs, they construct a causal graph of market drivers. This graph is then used to filter the inputs for quantitative models, ensuring the model relies on causal drivers (e.g., "Interest Rates") rather than spurious correlates (e.g., "Butter Production in Bangladesh").[30]

6. Domain Deep Dive: Cybersecurity and Social Network Analysis

These domains are characterized by adversarial interactions. The "theory" is the strategy of the attacker, which is hidden and changing.

6.1 Attack Graph Generation and Threat Attribution

Understanding the security posture of a network requires mapping all possible attack paths.

- **Knowledge Graph Neural Networks (KGNNs):** NeSy systems use KGNNs to model the complex, multi-stage nature of cyber attacks.[6] The KG captures the topology of the network and the vulnerabilities (CVEs).
- **Reasoning:** The system reasons over this graph to generate **Attack Graphs**. Unlike static scanners, these models can infer "multi-hop" exploits (e.g., "Attacker exploits Vuln A to gain User Access, then uses Credential Dumping to get Admin").
- **Attribution:** By extracting logical rules from the attack patterns, the system can perform **Threat Attribution**.[41] It matches the observed attack logic to the known TTPs (Tactics, Techniques, and Procedures) of threat actors (e.g., APT29), providing explainable evidence for forensic analysts.

6.2 Social Networks: Logic-Based Detection

In social networks, detecting bots or misinformation requires more than text analysis; it requires analyzing the *structure* of interaction.

- **Social Engineering Detection:** Hybrid models integrate neural classifiers (analyzing message content) with **Logic Probabilistic Models** (analyzing user context and relationship history).[42] A message might look benign to a specific NLP model, but the logic module flags it because "User A has never communicated with User B, and User B is external, and the request involves finance."
- **Fact-Checking:** NeSy systems verify claims by traversing a KG of verified facts. If an LLM generates a claim, the NeSy system checks for logical consistency against the KG (e.g., "Claim: X happened in 2024" vs KG: "X happened in 2022"). This **Fact-Checking via Reasoning** [43] is critical for mitigating the spread of misinformation in theory-poor social environments.

7. Synthesis: The Future of Interpretability and Trust

The convergence of Neural and Symbolic AI is not merely an academic exercise; it is a response to the practical demands of deploying AI in critical infrastructure. The transition from "Black Box" to "Glass Box" AI is essential for trust.

7.1 The "Green AI" Implication

An often-overlooked advantage of NeSy systems is **Data Efficiency**. By imposing symbolic constraints (priors), NeSy models restrict the search space of the neural network. This allows them to learn from significantly fewer examples than pure deep learning models. This aligns with the "Green AI" movement, suggesting that integrating domain knowledge is a path to reducing the massive carbon footprint of training foundation models.[44]

7.2 Open Challenges: Meta-Cognition and Scalability

Despite the progress, challenges remain.

- **Scalability:** Symbolic reasoning is computationally expensive (often NP-hard). Techniques like **Pruned Reasoning Paths** and **Efficient Rectification** [6] are being developed to make these systems viable for real-time applications, but trade-offs between depth of reasoning and latency persist.
- **Meta-Cognition:** A frontier for 2026 is **Meta-Cognition**—systems that not only reason but reason about their *own* reasoning.[45] Current systems can correct a prediction, but they lack the self-awareness to realize when their entire logical framework (KB) is obsolete and requires a fundamental paradigm shift.

8. Conclusion

The research landscape of 2024–2025 demonstrates that **Neuro-Symbolic AI** has matured into the definitive paradigm for knowledge discovery in **theory-poor, data-rich domains**. By fusing the perceptual power of Deep Learning with the reasoning rigor of Knowledge Graphs and Logic, these systems address the fundamental limitations of the previous decade of AI.

- **In Bioinformatics**, they are moving from predicting drug-disease links to reverse-engineering the logical circuits of the cell via **Abductive Learning** frameworks like ALIGNED and LogicSR.
- **In Finance**, they are replacing "profitability illusions" with auditable, logically sound **Trading Rules** and causal risk models derived from **Financial Knowledge Graphs**.
- **In Cybersecurity**, they are transforming threat detection from pattern matching to **Reasoning-Based Attribution** via Attack Graphs.
- **Methodologically**, the field has established robust foundations with **Graph Foundation Models** (GraphOracle) that reason inductively over unseen entities, and formalized **Interpretability Metrics** (Fidelity, Sparsity) that make AI auditability a quantitative science.

As we look forward, the distinction between "learning from data" and "reasoning from theory" is dissolving. The Neuro-Symbolic systems of the future will be **Automated Scientists**: observing the world through neural sensors, formulating logical hypotheses, testing them against data, and refining their internal theory of the world—unlocking the latent structure of our most complex, chaotic, and data-rich environments.

Works cited

1. Adaptive Data-Knowledge Alignment in Genetic Perturbation Prediction - arXiv, accessed January 8, 2026, https://arxiv.org/html/2510.00512v1
2. LogicSR: prior-guided symbolic regression for gene regulatory network inference from single-cell transcriptomics data | Briefings in Bioinformatics | Oxford Academic, accessed January 8, 2026, https://academic.oup.com/bib/article/26/6/bbaf621/8339795
3. Root Cause Analysis of Cast Product Defects with Two-Branch Reasoning

Network Based on Continuous Casting Quality Knowledge Graph - MDPI, accessed January 8, 2026, https://www.mdpi.com/2076-3417/15/13/6996
4. AutoQuant: An Auditable Expert-System Framework for Execution-Constrained Auto-Tuning in Cryptocurrency Perpetual Futures - arXiv, accessed January 8, 2026, https://arxiv.org/pdf/2512.22476
5. INFERRING THE INVISIBLE: RECURRENT NEURO- SYMBOLIC FORWARD CHAINING NETWORK - OpenReview, accessed January 8, 2026, https://openreview.net/pdf/4973514e568af585277e301012142c787eb1c155.pdf
6. Neuro-Symbolic AI for Cybersecurity: State of the Art, Challenges, and Opportunities - arXiv, accessed January 8, 2026, https://arxiv.org/html/2509.06921v1
7. Neuro-Symbolic AI: A Foundational Analysis of the Third Wave's Hybrid Core, accessed January 8, 2026, https://gregrobison.medium.com/neuro-symbolic-ai-a-foundational-analysis-of-the-third-waves-hybrid-core-cc95bc69d6fa
8. GraphOracle: Efficient Fully-Inductive Knowledge Graph Reasoning via Relation-Dependency Graphs - arXiv, accessed January 8, 2026, https://arxiv.org/html/2505.11125v2
9. A Neuro-Symbolic System over Knowledge Graphs for Link Prediction, accessed January 8, 2026, https://www.semantic-web-journal.net/content/neuro-symbolic-system-over-knowledge-graphs-link-prediction
10. Neurosymbolic AI for reasoning over knowledge graphs: A survey, accessed January 8, 2026, https://www.research.ed.ac.uk/en/publications/neurosymbolic-ai-for-reasoning-over-knowledge-graphs-a-survey/
11. AI Reasoning in Deep Learning Era: From Symbolic AI to Neural–Symbolic AI - MDPI, accessed January 8, 2026, https://www.mdpi.com/2227-7390/13/11/1707
12. [2505.11125] GraphOracle: A Foundation Model for Knowledge Graph Reasoning - arXiv, accessed January 8, 2026, https://arxiv.org/abs/2505.11125
13. A universal approach for extracting interpretable neuro-symbols - SPIE Digital Library, accessed January 8, 2026, https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13681/136811Q/A-universal-approach-for-extracting-interpretable-neuro-symbols/10.1117/12.3074791.full
14. Neurosymbolic Methods for Explainable Graph Neural Networks: A Survey - CEUR-WS.org, accessed January 8, 2026, https://ceur-ws.org/Vol-4125/paper_4.pdf
15. LLM-empowered knowledge graph construction: A survey - arXiv, accessed January 8, 2026, https://arxiv.org/html/2510.20345v1
16. Schema Induction & Logical Representation - Emergent Mind, accessed January 8, 2026, https://www.emergentmind.com/topics/schema-induction-and-logical-representation
17. Neural-Symbolic Reasoning over Knowledge Graphs: A Survey from a Query

Perspective - SIGKDD, accessed January 8, 2026, https://kdd.org/exploration_files/p124-Neural_Symbolic_KGR_survey.pdf
18. Neuro-Symbolic AI in 2024: A systematic review - Brandon Colelough, accessed January 8, 2026, https://brandoncolelough.com/Files/LNSAI_survey_paper.pdf
19. The Future Is Neuro-Symbolic: Where Has It Been, and Where Is It Going? - Rivista AI, accessed January 8, 2026, https://www.rivista.ai/wp-content/uploads/2025/11/Belle_Marcus_AAAI-2.pdf
20. Explainable Ai for Knowledge Graph-Based Drug Repurposing: Methods, Challenges, and Interpretability Frameworks - IEEE Computer Society, accessed January 8, 2026, https://www.computer.org/csdl/proceedings-article/icebe/2025/903800a015/2cED3QdwCZi
21. A review of Neuro-Symbolic AI integrating reasoning and learning for advanced cognitive systems | Request PDF - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/392044200_A_review_of_Neuro-Symbolic_AI_integrating_reasoning_and_learning_for_advanced_cognitive_systems
22. Related papers: GraphOracle: A Foundation Model for Knowledge Graph Reasoning - Fugu Machine Translator, accessed January 8, 2026, https://fugumt.com/fugumt/paper_check/2505.11125v1_enmode
23. Advancing Abductive Reasoning in Knowledge Graphs through Complex Logical Hypothesis Generation | Request PDF - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/384209454_Advancing_Abductive_Reasoning_in_Knowledge_Graphs_through_Complex_Logical_Hypothesis_Generation
24. Diagnosing and Addressing Pitfalls in KG-RAG Datasets: Toward More Reliable Benchmarking - arXiv, accessed January 8, 2026, https://arxiv.org/html/2505.23495v3
25. Recent Applications of Explainable AI (XAI): A Systematic Literature Review - MDPI, accessed January 8, 2026, https://www.mdpi.com/2076-3417/14/19/8884
26. Improving Interpretability and Accuracy in Neuro-Symbolic Rule Extraction Using Class-Specific Sparse Filters - arXiv, accessed January 8, 2026, https://arxiv.org/html/2501.16677v1
27. Adaptive Data-Knowledge Alignment in Genetic Perturbation Prediction - arXiv, accessed January 8, 2026, https://arxiv.org/html/2510.00512
28. CausalTrace: A Neurosymbolic Causal Analysis Agent for Smart Manufacturing - arXiv, accessed January 8, 2026, https://arxiv.org/html/2510.12033v1
29. Large Causal Models from Large Language ModelsDraft under revision. - arXiv, accessed January 8, 2026, https://arxiv.org/html/2512.07796v1
30. Integrating causal discovery and machine learning for enhanced financial forecasting - TU Wien's reposiTUm, accessed January 8, 2026, https://repositum.tuwien.at/bitstream/20.500.12708/216394/1/Schwarzinger%20Pia%20-%202025%20-%20Integrating%20causal%20discovery%20and%20machine%20learning%20for...pdf
31. Neural Description Logic Reasoning over Incomplete Knowledge Bases - OpenReview, accessed January 8, 2026,

https://openreview.net/forum?id=4qRCiEZGKd
32. Neural Reasoning for Robust Instance Retrieval in $\mathscr{SHOIQ}$ - arXiv, accessed January 8, 2026, https://arxiv.org/html/2510.20457v1
33. Neurosymbolic Learning for Predicting Cell Fate Decisions from Longitudinal Single-Cell Transcriptomics in Paediatric Acute Myeloid Leukemia - arXiv, accessed January 8, 2026, https://arxiv.org/html/2508.13199v1
34. Adaptive Data-Knowledge Alignment in Genetic Perturbation Prediction - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/396095220_Adaptive_Data-Knowledge_Alignment_in_Genetic_Perturbation_Prediction
35. A comprehensive large scale biomedical knowledge graph for AI powered data driven biomedical research - PMC - NIH, accessed January 8, 2026, https://pmc.ncbi.nlm.nih.gov/articles/PMC10760044/
36. Knowledge-graph-based explainable AI: A systematic review - PMC - NIH, accessed January 8, 2026, https://pmc.ncbi.nlm.nih.gov/articles/PMC11316662/
37. Enhanced decision making mechanism of rule-based genetic network programming for creating stock trading signals | Request PDF - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/257405004_Enhanced_decision_making_mechanism_of_rule-based_genetic_network_programming_for_creating_stock_trading_signals
38. Innovative Portfolio Optimization Using Deep Q-Network Reinforcement Learning | Request PDF - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/390751224_Innovative_Portfolio_Optimization_Using_Deep_Q-Network_Reinforcement_Learning
39. Template-based Explainable Inference over High-Stakes Financial Knowledge Graphs - OpenProceedings.org, accessed January 8, 2026, https://openproceedings.org/2025/conf/edbt/paper-135.pdf
40. (PDF) AI-AUGMENTED CYBERSECURITY: GRAPH NEURAL NETWORKS FOR PREDICTING NATION-STATE CYBERATTACKS - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/396109613_AI-AUGMENTED_CYBERSECURITY_GRAPH_NEURAL_NETWORKS_FOR_PREDICTING_NATION-STATE_CYBERATTACKS
41. (PDF) Neuro-Symbolic Cybersecurity Models for Explainable Threat Attribution and Attack Surface Governance - ResearchGate, accessed January 8, 2026, https://www.researchgate.net/publication/397899149_Neuro-Symbolic_Cybersecurity_Models_for_Explainable_Threat_Attribution_and_Attack_Surface_Governance
42. Neuro- Symbolic AI and Behavioral Intelligence: - IGI Global, accessed January 8, 2026, https://www.igi-global.com/ViewTitle.aspx?TitleId=398514&isxn=9798337350127
43. Neuro-Symbolic Artificial Intelligence: A Task-Directed Survey in the Black-Box Models Era - IJCAI, accessed January 8, 2026, https://www.ijcai.org/proceedings/2025/1157.pdf

44. Neurosymbolic AI as an antithesis to scaling laws | PNAS Nexus - Oxford Academic, accessed January 8, 2026, https://academic.oup.com/pnasnexus/article/4/5/pgaf117/8134151
45. Neuro-Symbolic AI in 2024: A Systematic Review - arXiv, accessed January 8, 2026, https://arxiv.org/html/2501.05435v1