Regulation-Aware Neuro-Symbolic Legal World Models for Cross-Border Commerce: A Position Paper

David Scott Lewis
AIXC
Zaragoza, Spain
research@aiexecutiveconsulting.commane

Abstract

This position paper proposes a research agenda at the intersection of Artificial Intelligence and Law (AI-LAW), neuro-symbolic AI (NeSy), and world models, focusing on the complex demands of international commercial practice. Cross-border transactions, AI-enabled supply chains, and arbitration increasingly rely on opaque foundation models, while emerging regulatory regimes, notably the EU AI Act, impose stringent risk-based obligations on high-risk AI systems. Purely neural models struggle to provide the explainability, reliability, and regulatory alignment required in these high-stakes environments. We argue that hybrid neuro-symbolic architectures and simulation-based world models offer a concrete pathway to more trustworthy legal AI. We introduce the concept of Regulation-Aware Neuro-Symbolic Legal World Models (R-NSLWMs): systems that integrate (i) extensive legal, regulatory, and contractual knowledge graphs aligned with benchmarks like LegalBench; (ii) domain-specific commercial world models (Legal Digital Twins) that simulate contractual performance and regulatory environments; and (iii) neuro-symbolic reasoning layers that enforce legal and physical constraints while generating human-legible explanations. We outline a four-layer architecture for R-NSLWMs, discuss use cases in AI-regulation-aware contracting and arbitration support, and propose a roadmap for developing the necessary benchmarks and methodologies. This agenda aims to bridge the gap between abstract regulatory principles and concrete technical implementations in global commerce.

1 Introduction: The Need for Structure in Legal AI

The integration of foundation models, particularly Large Language Models (LLMs), into legal practice marks a transformative shift. LLMs are rapidly being deployed in legal research platforms, contract drafting tools, and dispute resolution systems (Choi, Monahan, and Schwarcz 2024; Lai et al. 2023). This adoption promises significant productivity gains and the potential to narrow access-to-justice gaps (Chien et al. 2025). However, the reliance on purely neural, opaque models introduces substantial risks. Hallucinations, embedded biases, and the inability to provide structured, verifiable explanations are critical shortcomings in high-stakes legal environments (Bommasani et al. 2021; Regalia 2024).

In international commercial law, these risks are amplified. AI systems are integral to complex, cross-border operations, including sanctions screening, Anti-Money Laundering (AML), credit scoring, and supply-chain optimization. The reliability and explainability of these systems directly impact enforceability, liability, and regulatory exposure (Wagner 2023). Furthermore, the emergence of comprehensive regulatory frameworks, most notably the EU AI Act (Regulation (EU) 2024/1689), mandates a risk-based approach, imposing strict obligations on providers and deployers of high-risk AI systems (RAND Corporation 2024; Almada 2025). Ensuring compliance requires AI systems that are not merely accurate on average, but demonstrably trustworthy, fair, and transparent (Alsagheer et al. 2025).

The limitations of current LLMs suggest that "trustworthy AI" cannot be achieved through scaling laws alone (Kambhampati 2022). This realization has fueled the resurgence of neuro-symbolic AI (NeSy), which seeks to integrate the robust learning capabilities of neural networks with the structured reasoning and interpretability of symbolic methods (Garcez et al. 2022). NeSy approaches aim to make foundation models more deterministic, interpretable, and reliable by incorporating explicit knowledge graphs (KGs), logical rules, and constraints (DeLong et al. 2023; Cunnington et al. 2024).

Concurrently, the field of world models has matured significantly. Systems like Genie (Bruce et al. 2024) and DreamerV3 (Hafner et al. 2023) demonstrate the ability to learn complex, interactive environments from raw data, enabling simulation and planning in virtual spaces. In domains like AI-for-Science, these models serve as "digital laboratories," integrating multimodal data with scientific KGs to simulate experiments and enforce physical constraints (Rueden et al. 2023; Guo et al. 2024).

Our position is that international commercial law demands a synthesis of these trends. This synthesis aligns with the goals of initiatives like the AAAI AI-LAW Bridge (AIL 2025) and the NeusymBridge workshop (Neu 2025). We propose the development of **Regulation-Aware Neuro-Symbolic Legal World Models (R-NSLWMs)**. These sys-

AAAI Acknowledgement: This material is prepared for submission to the AAAI 2026 conference.

Proceedings of Bridge between Artificial Intelligence and Law 2026 (Short papers), pages 271–275
Januray 21, 2026, AAAI 2026 Bridge Program, Singapore

tems move beyond treating law as a textual corpus for LLMs. Instead, R-NSLWMs embed legal and regulatory structures directly into hybrid AI architectures. They combine (1) comprehensive legal KGs; (2) commercial world models that simulate the dynamics of international trade and finance; and (3) neuro-symbolic reasoning layers that enforce constraints and generate legally intelligible explanations.

This paper outlines a vision and research agenda for R-NSLWMs. We articulate a four-layer architecture, explore key use cases, and identify critical research challenges. By grounding legal reasoning in structured knowledge and simulation, R-NSLWMs offer a concrete pathway toward AI systems that can navigate the complexities of global commerce while adhering to evolving regulatory mandates.

2 Background and Converging Trends

The proposed R-NSLWM framework draws upon three distinct but converging research areas: Trustworthy AI in legal contexts, Neuro-Symbolic AI and Knowledge Graphs, and Foundation World Models.

2.1 Trustworthy AI and Legal Benchmarks

The discourse surrounding AI in law has increasingly centered on trustworthiness, encompassing explainability, fairness, accountability, and regulatory compliance (Balcioğlu, Çelik, and Altindağ 2025). This focus is driven by the high-stakes nature of legal decisions and the potential for AI-driven errors or biases to cause significant harm (Wagner 2023).

A critical component of evaluating trustworthiness is the development of specialized benchmarks. Initiatives like LegalBench (Guha et al. 2023) and LexGLUE (Chalkidis et al. 2022) provide standardized tasks for measuring legal reasoning capabilities in LLMs, covering areas such as statutory interpretation and contract analysis. These benchmarks highlight that while LLMs excel at information retrieval, they often struggle with complex, multistep legal reasoning and consistency (Lai et al. 2023).

Furthermore, the regulatory landscape is rapidly evolving. The EU AI Act represents the most comprehensive attempt to regulate AI based on risk (RAND Corporation 2024). For high-risk systems, the Act mandates stringent requirements for data governance, documentation, human oversight, and robustness. Compliance necessitates systems capable of providing transparent explanations and audit trails, requirements that purely data-driven models struggle to meet (Almada 2025).

2.2 Neuro-Symbolic AI and Knowledge Graphs

Neuro-Symbolic AI (NeSy) addresses the limitations of purely neural models by combining deep learning with symbolic knowledge representation and reasoning (Garcez et al. 2022). The goal is to create hybrid systems that are more interpretable, data-efficient, and robust (Kambhampati 2022).

Knowledge Graphs (KGs) are central to many NeSy architectures. KGs provide a structured representation of entities and their relationships, encoding domain knowledge in a machine-readable format (Hogan et al. 2021). In NeSy systems, KGs can be used to constrain the output of neural models, inject domain knowledge during training, or provide a symbolic structure for interpreting the model's reasoning process (DeLong et al. 2023).

Recent work explores various integration strategies. These include differentiable reasoning frameworks like Logic Tensor Networks (LTNs) (Badreddine et al. 2020), neural theorem provers, and architectures that distill symbolic rules from LLMs (Cunnington et al. 2024). In high-stakes domains such as medicine and AI-for-Science, hybrid models integrating KGs with foundation models are already being used to enhance interpretability and enforce constraints (Vidal et al. 2025; Rueden et al. 2023). For instance, in drug discovery, knowledge-augmented graph machine learning demonstrates how domain KGs can guide the exploration of molecular structures (Perdomo-Quinteiro et al. 2024; Ding et al. 2025).

The structured nature of legal and regulatory knowledge makes it a prime candidate for NeSy approaches. Representing laws and contracts as KGs can enable more precise and verifiable legal reasoning than relying solely on the latent representations of LLMs.

2.3 World Models and Legal Digital Twins

World models learn representations of environments and their dynamics, enabling agents to simulate future states and plan actions (Hafner et al. 2023). The emergence of foundation world models, such as Genie (Bruce et al. 2024), demonstrates the capability to generate complex, interactive virtual environments from unlabeled data. Commercial platforms like Marble are also emerging for spatial AI applications (World Labs 2025).

In AI-for-Science, world models are utilized as digital laboratories to simulate complex systems, often integrating multimodal data and domain-specific KGs (Guo et al. 2024; López et al. 2025).

We propose the analogous concept of a **Legal Digital Twin**: a world model of a commercial ecosystem (e.g., a supply chain, financial network, or infrastructure project) that is constrained not only by physical and economic dynamics but also by legal and regulatory rules. An R-NSLWM integrates the legal KG directly into the simulation environment, allowing for the real-time evaluation of contractual performance and regulatory compliance under varying conditions.

3 The R-NSLWM Architecture

We define a Regulation-Aware Neuro-Symbolic Legal World Model (R-NSLWM) as a hybrid system designed to model, simulate, and reason about international commercial activities within a multi-jurisdictional regulatory framework. The architecture consists of four interlocking layers.

3.1 Layer 1: Legal, Regulatory, and Contractual Knowledge Graphs

The foundational layer is a comprehensive, multi-jurisdictional Legal Knowledge Graph (LKG). This

KG moves beyond simple ontologies to represent the complex structure of international law and commerce.

Structure and Content Nodes in the LKG represent regulatory provisions (e.g., EU AI Act articles), international law (e.g., CISG, UNCITRAL model laws), contractual elements (clauses, warranties), and jurisprudence (key arbitral awards).

Edges encode precise relationships, such as "implements," "derogates from," or "is interpreted by." Attributes capture critical metadata, including jurisdiction, temporal validity, risk classification, and roles (e.g., AI provider vs. deployer).

Alignment and Reasoning The LKG is explicitly aligned with legal benchmarks (Guha et al. 2023; Chalkidis et al. 2022). Benchmark tasks are reformulated as KG queries or inference problems. Technically, this layer leverages advancements in KG integration and neuro-symbolic reasoning (Hogan et al. 2021). Techniques such as graph neural networks with logical constraints can be adapted to reason over the LKG (DeLong et al. 2023). Recent work on causal AI in legal language processing suggests methods for representing causal relationships (e.g., actions leading to regulatory violations) within the graph (Tritto and Ponce 2025).

3.2 Layer 2: Commercial and Infrastructural World Models

The second layer consists of domain-specific world models (Legal Digital Twins) that simulate the commercial and infrastructural systems relevant to international trade. These models are inspired by foundation world models (Bruce et al. 2024; Hafner et al. 2023) but are tailored to represent contractual performance, operational events, and regulatory states.

Examples include models of global supply chain dynamics (simulating delays and disruptions) and financial ecosystems (simulating credit events and macroeconomic scenarios).

These world models integrate multimodal data and leverage techniques from AI-for-Science, where KGs are used to guide representation learning and ensure consistency with domain knowledge (Ding et al. 2025). Crucially, the states within the world model are annotated with pointers to the LKG. A simulated logistics delay might link to relevant contractual clauses (e.g., liquidated damages) or regulatory obligations.

3.3 Layer 3: Neuro-Symbolic Constraint and Explanation Layer

This layer connects the LKG (Layer 1) and the World Models (Layer 2) via neuro-symbolic constraints. It treats legal and physical rules as explicit predicates over world-model trajectories and agent policies.

Constraint Enforcement Neuro-symbolic AI provides the tools to encode and enforce these constraints during simulation and planning (Garcez et al. 2022). This includes legal constraints (e.g., banned AI practices, mandatory procedural steps, fairness metrics (Alsagheer et al. 2025)) and operational constraints (e.g., performance thresholds).

Frameworks like Logic Tensor Networks (Badreddine et al. 2020) or differentiable rule systems can be used to define constraint losses or logical satisfaction objectives. This ensures that simulations remain faithful to both the physical reality and the legal framework. Causal AI methods can be employed to trace how specific actions lead to regulatory violations within the simulation (Tritto and Ponce 2025).

Legally Intelligible Explanations Because the constraints are symbolic and grounded in the LKG, the system can generate structured, legally intelligible explanations. For any simulated scenario, the R-NSLWM can articulate *why* a particular outcome occurred, referencing both the world-model dynamics and the triggering legal rules. This directly addresses the explainability requirements crucial for regulatory compliance and dispute resolution (Vidal et al. 2025).

3.4 Layer 4: Agentic Workflows for Commercial Practice

The final layer instantiates concrete workflows where human experts interact with neuro-symbolic agents operating on top of the R-NSLWM.

Examples include:

- **Regulation-Aware Drafting:** Agents assist in drafting AI-related contractual clauses by simulating performance under different regulatory scenarios (Layer 2) and ensuring compliance with the LKG (Layer 1).
- **Benchmark-Aligned Evaluation:** Agents use standardized benchmarks to evaluate AI systems used in a transaction, mapping performance to regulatory obligations.
- **Dispute Resolution Support:** Agents assist arbitrators in evaluating counterfactual scenarios within the world model, providing evidence grounded in both factual simulation and legal constraints.

In these workflows, the agent explicitly surfaces its reasoning, allowing human experts to validate or override the outputs, ensuring meaningful human oversight (RAND Corporation 2024).

4 Use Cases in Cross-Border Commerce

The R-NSLWM framework offers significant potential across international commercial practice. We highlight three key use cases.

4.1 AI-Regulation KG for Transnational Deals

The complexity of navigating AI regulations across multiple jurisdictions poses a significant challenge for transnational deals. The EU AI Act, with its extraterritorial reach, exemplifies this challenge (Almada 2025).

An R-NSLWM can support this by utilizing a specialized AI-Regulation KG. This KG would integrate the EU AI Act, national statutes, and sector-specific guidelines, encoding risk categories, obligations, and timelines (RAND Corporation 2024).

Practitioners can leverage this system for deal structuring and due diligence (automatically classifying AI systems and identifying obligations) and clause generation (drafting warranties and covenants that reflect regulatory requirements). The system can validate proposed clauses against the KG to ensure consistency. Furthermore, the world model layer can simulate the impact of anticipated regulatory changes on existing contracts, allowing for proactive adaptation.

4.2 Explainable Risk Modeling in Trade Finance

Financial institutions heavily rely on AI for credit scoring, AML, and portfolio risk modeling. Many of these applications may be classified as high-risk, requiring high levels of transparency and fairness (Balcioğlu, Çelik, and Altindağ 2025).

An R-NSLWM for trade finance would combine a world model of credit portfolios with a KG of financial regulations across jurisdictions. The neuro-symbolic layer would enforce constraints related to fairness metrics and regulatory triggers.

The system can simulate credit decisions under various scenarios, flagging instances where constraints are violated. Crucially, it generates human-readable explanations for portfolio-level risk exposures, grounded in both the simulation dynamics and the underlying legal rules. These explanations are essential for internal risk management and regulatory reporting (Vidal et al. 2025).

4.3 World-Model-Based Arbitration Support

International arbitration frequently involves complex factual disputes concerning contractual performance, causation, and damages. Evaluating these disputes requires sophisticated counterfactual reasoning.

The R-NSLWM provides a powerful tool for supporting arbitration through Legal Digital Twins. By simulating the relevant commercial environment (e.g., a logistics chain) constrained by the contractual terms and applicable laws, the system allows counsel and tribunals to explore "what if" scenarios.

For example, the system can evaluate the impact of different clause interpretations on liability allocation. Because the simulations are generated by a neuro-symbolic system, the results are accompanied by explanations that link the simulated events to specific legal provisions. This enhances the transparency and rigor of the arbitration process (Tritto and Ponce 2025). We also propose the development of new benchmark tasks tailored to arbitration workflows, extending the scope of existing benchmarks (Guha et al. 2023).

5 Research Agenda and Challenges

The development of R-NSLWMs presents a rich interdisciplinary research agenda, requiring addressing significant challenges in representation, inference, and evaluation.

5.1 Representation and Semantics

A fundamental challenge lies in the representation of legal norms within KGs and symbolic constraint languages. Legal concepts are often open-textured (e.g., "reasonableness," "good faith"), resisting straightforward formalization (Lai et al. 2023). Research is needed to develop representation techniques that capture the nuance of legal language while remaining computationally tractable.

Furthermore, aligning the semantics of legal norms (discrete, symbolic) with the continuous representations of world models is non-trivial. This requires developing abstractions that map raw world-model trajectories (e.g., event logs) to legally relevant facts (e.g., breach, causation). Techniques from symbolic scene-graph extraction may offer starting points (Hafner et al. 2023).

5.2 Neuro-Symbolic Inference and Learning

Selecting and adapting neuro-symbolic frameworks for the legal domain is a key research area. Candidates include Logic Tensor Networks (Badreddine et al. 2020) and causal NeSy architectures (Rueden et al. 2023). The challenge is to train these hybrid models such that they remain faithful to explicit legal constraints while generalizing effectively from data. Multi-objective training strategies combining predictive accuracy and constraint satisfaction will be crucial.

Moreover, exploiting causal AI methods to draw robust causal conclusions from world-model simulations is essential, particularly when using simulated evidence in arbitral or regulatory contexts (Tritto and Ponce 2025).

5.3 Benchmarks, Evaluation, and Governance

To drive progress, we need benchmarks that specifically capture "regulation-aware legal reasoning" in cross-border commerce. Existing benchmarks (Guha et al. 2023; Chalkidis et al. 2022) must be extended to include tasks related to AI regulation and counterfactual reasoning in world models.

Evaluating the explanations generated by R-NSLWMs is also critical. Metrics must assess legal soundness, completeness, and usefulness to human decision-makers. This requires empirical studies on human-AI interaction in legal contexts (Chien et al. 2025).

Finally, the governance of the Legal KGs and world models themselves poses significant challenges. Ensuring the integrity, currency, and fairness of these foundational components is paramount to the trustworthiness of the overall system (Alsagheer et al. 2025).

6 Conclusion

The increasing complexity of international commerce and the stringent requirements of emerging AI regulations demand a paradigm shift in legal AI. Purely neural approaches lack the structural guarantees, explainability, and regulatory alignment necessary for high-stakes cross-border environments. We have proposed Regulation-Aware Neuro-Symbolic Legal World Models (R-NSLWMs) as a pathway to trustworthy legal AI. By integrating comprehensive legal knowledge graphs, dynamic commercial world models, and neuro-symbolic reasoning, R-NSLWMs provide a framework for modeling, simulating, and explaining complex commercial activities within their legal context, while enhancing compliance, transparency, and the rule of law.

References

2025. AI-LAW @ AAAI 2026 – Bridge between Artificial Intelligence and Law. AI-LAW Bridge website.

2025. NeusymBridge Workshop @ AAAI 2026. Neusym-Bridge website.

Almada, M. 2025. Will the EU AI Act Shape Global Regulation? *Network Law Review*.

Alsagheer, D. R.; et al. 2025. Consistency and Fairness as Keys to Reliable AI. In *Proceedings of the Association for Computational Linguistics (ACL)*.

Badreddine, S.; d'Avila Garcez, A. S.; Serafini, L.; and Spranger, M. 2020. Logic Tensor Networks. *Artif. Intell.*, 303: 103649.

Balcioğlu, Y. S.; Çelik, A. A.; and Altindağ, E. 2025. A turning point in AI: Europe's human-centric approach to technology regulation. *Journal of Responsible Technology*, 100128.

Bommasani, R.; et al. 2021. On the Opportunities and Risks of Foundation Models. arXiv:2108.07258.

Bruce, J.; et al. 2024. Genie: Generative Interactive Environments. arXiv:2402.15391.

Chalkidis, I.; et al. 2022. LexGLUE: A Benchmark Dataset for Legal Language Understanding in English. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Chien, C. V.; et al. 2025. Generative AI and Legal Aid: Results from a Field Study and Design Agenda. *Loyola of Los Angeles Law Review*.

Choi, J. H.; Monahan, A.; and Schwarcz, S. L. 2024. Lawyering in the Age of Artificial Intelligence. *Minnesota Law Review*.

Cunnington, D.; et al. 2024. The Role of Foundation Models in Neuro-Symbolic Learning and Reasoning. In *Neural-Symbolic Learning and Reasoning*.

DeLong, L. N.; et al. 2023. Neurosymbolic AI for Reasoning over Knowledge Graphs: A Survey. arXiv:2302.07200.

Ding, K.; et al. 2025. Bridging Data and Discovery: A Survey on Knowledge Graphs in AI for Science.

Garcez, A. d.; et al. 2022. *Neurosymbolic Artificial Intelligence: The State of the Art*. IOS Press.

Guha, N.; et al. 2023. LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models. In *NeurIPS 2023 Datasets and Benchmarks Track*.

Guo, S.; Shariatmadari, A. H.; Xiong, G.; and Zhang, A. 2024. Embracing Foundation Models for Advancing Scientific Discovery. *2024 IEEE International Conference on Big Data (BigData)*, 1746–1755.

Hafner, D.; et al. 2023. Mastering Diverse Domains through World Models. arXiv:2301.04104.

Hogan, A.; et al. 2021. Knowledge Graphs. *ACM Computing Surveys (CSUR)*, 54(4).

Kambhampati, S. 2022. The case for hybrid AI: Combining symbolic and neural approaches. *AI Magazine*.

Lai, J.; Gan, W.; Wu, J.; Qi, Z.; and Yu, P. S. 2023. Large Language Models in Law: A Survey. *ArXiv*, abs/2312.03718.

López, V.; et al. 2025. Enhancing Foundation Models for Scientific Discovery via Multimodal Knowledge Graph Representations. *Journal of Web Semantics*.

Perdomo-Quinteiro, P.; et al. 2024. Knowledge Graphs and Explainable AI for Drug Repurposing on Rare Diseases. *Neurosymbolic Artificial Intelligence*.

RAND Corporation. 2024. Risk-Based AI Regulation: A Primer on the Artificial Intelligence Act of the European Union. Technical Report RRA3243-3, RAND Corporation.

Regalia, J. 2024. How Generative AI Is Transforming Legal Writing and Research. *UNLV Faculty Publications*.

Rueden, L. v.; et al. 2023. Neuro-symbolic AI in the sciences: A survey. *Computing in Science & Engineering*.

Tritto, A.; and Ponce, Á. 2025. Causal Artificial Intelligence in Legal Language Processing: A Systematic Review. *Entropy*.

Vidal, M.-E.; et al. 2025. Integrating Knowledge Graphs with Symbolic AI: The Path to Interpretable Hybrid AI Systems in Medicine. *Journal of Web Semantics*.

Wagner, G. 2023. Liability for AI-related harm: A comparative overview. *Journal of European Tort Law*.

World Labs. 2025. Marble: From Pixels to Worlds. World Labs website.