# MERL-T: a Multi-Expert architecture for trustworthy Artificial Legal Intelligence

**Daniele Allega1\***, **Guglielmo Puzio2**

1 Mercatorum University, Piazza Mattei, 10, 00186 Roma RM, Italy

 [https://orcid.org/0009-0006-5359-6102](https://orcid.org/0009-0006-5359-6102%20)

e-mail: [dallega@luiss.it](mailto:dallega@luiss.it), [daniele.allega@studenti.unimercatorum.it](mailto:daniele.allega@studenti.unimercatorum.it)

2 LUISS “Guido Carli” University, Viale Romania, 32, 00197 Roma RM, Italy

 <https://orcid.org/0009-0001-4366-2632>

e-mail: [guglielmo.puzio@studenti.luiss.it](mailto:guglielmo.puzio@studenti.luiss.it)

\* Corresponding author

**Abstract**: *Contemporary Large Language Models fail to capture the fundamental epistemic structure of legal reasoning. Law operates through a constitutive duality: abstract principles providing interpretive scaffolding and concrete rules governing specific situations. Generic LLMs collapse this duality into undifferentiated text, producing plausible but epistemically unfounded outputs. We introduce MERL-T (Multi-Expert Retrieval Legal Transformer), a five-layer pipeline architecture operationalizing legal epistemology through specialized components that preserve the heterogeneity of legal reasoning itself. MERL-T implements epistemic fidelity through: (1) a knowledge graph (KG) encoding legal relationships as explicit, navigable structures rather than implicit weights, distinguishing normative, conceptual, jurisprudential, and doctrinal entities with temporal versioning; (2) orchestrated retrieval agents integrating heterogeneous sources (structured norms, case law, doctrine) through LLM-based dynamic planning; (3) four expert modules grounded in distinct legal philosophies—literal interpretation (legal positivism), systemic-teleological reasoning (legal teleology), principles balancing (constitutional principlism), and precedent analysis (legal realism)—each maintaining internal coherence according to its epistemic commitments; (4) a synthesis layer employing convergent mode when experts agree on conclusions but reason differently, and divergent mode when disagreement reflects genuine legal ambiguity, refusing to force false consensus. Unlike monolithic approaches that collapse methodological plurality into single outputs, MERL-T preserves epistemic plurality by maintaining multiple valid interpretations when legal materials genuinely support divergent readings. Our architectural analysis demonstrates how each design choice follows necessarily from the epistemic requirements of legal reasoning itself, providing a model for trustworthy AI in specialized domains where reasoning structure matters as much as linguistic fluency.*

**Keywords:** Legal Epistemology, Multi-Expert Systems, Knowledge Graphs, Legal Reasoning, Artificial Legal Intelligence

**JEL Classification:** *C45; C55; K00; K40; O31.*

**Note:** *the full version of this abstract is available in the references (Allega & Puzio, 2025b)*

# 1. Introduction

Legal reasoning is specialized practical rationality operating through distinctive epistemic structures (Hart, 1961): abstract principles (constitutional values, fundamental rights) providing conceptual scaffolding, and concrete rules (statutory provisions, regulatory requirements) governing particular situations (Dworkin, 1977). Competent legal reasoning requires navigating between these levels through multiple interpretive methodologies. Contemporary LLMs, trained on undifferentiated corpora, fail to respect this epistemic structure, producing outputs that are linguistically fluent but epistemically unfounded through: (1) epistemic opacity - cannot explain why interpretations are legally valid, (2) structural blindness - miss hierarchical relationships structuring legal knowledge, and (3) methodological monism - collapse different interpretive approaches into single undifferentiated output.

If the problem is epistemic, the solution must be architectural: build epistemic structure into system design. MERL-T (Allega & Puzio, 2025b) operationalizes legal epistemology through specialized components that (1) represent legal knowledge structure explicitly (Allega & Puzio, 2025a), (2) reason according to distinct legal methodologies, and (3) preserve epistemic plurality when disagreement reflects genuine ambiguity (Allega & Puzio, 2025c). This has three architectural consequences: *Explicit Knowledge Representation* (structured knowledge graph (KG) capturing relationships between concepts, norms, principles, precedents), *Methodological Specialization* (distributed across specialized expert modules grounded in distinct legal philosophies), and *Uncertainty Preservation* (preserve disagreement when multiple interpretations are valid). MERL-T implements these through five layers: *Preprocessing* (query understanding + KG enrichment), *Orchestration* (LLM-based routing + retrieval agents), *Reasoning* (four specialized experts + synthesis), *Storage* (multi-database architecture), and *Learning* (community feedback with uncertainty preservation) (Allega & Puzio, 2025c).

# 2. Literature Review on the Five-layer architecture

MERL-T processes Italian legal queries through five interconnected layers:

1. **Preprocessing Layer:** six-stage adaptive processing: abbreviation expansion, entity extraction, concept mapping, intent classification, complexity estimation, temporal detection. KG Enrichment Engine maps concepts to norms/principles/precedents through graph traversal. (Hamilton et al., 2017)
2. **Orchestration Layer:** LLM-based Router generates dynamic execution plans selecting retrieval agents (KG Agent, API Agent, VectorDB Agent) and reasoning experts based on query characteristics (Brown et al., 2020; Anthropic, 2024).
3. **Reasoning Layer:** four specialized experts (Literal Interpreter, Systemic-Teleological Reasoner, Principles Balancer, Precedent Analyst) perform legal analysis grounded in distinct epistemological frameworks. Synthesizer combines outputs through convergent or divergent synthesis. (Kipf & Welling, 2017)
4. **Storage Layer:** Neo4j KG for structured relationships, ChromaDB for semantic search, PostgreSQL for metadata. Data ingestion pipeline processes legislation, jurisprudence, doctrine. (Bordes et al., 2013; Lehmann et al., 2015).
5. **Learning Layer:** community feedback with dynamic authority weighting enables progressive improvement while maintaining epistemic fidelity (Allega, 2025; Allega & Puzio 2025c).

# 3. Knowledge graph: encoding legal epistemology

The KG is an epistemic structure making implicit legal relationships explicit and navigable, serving three functions: conceptual scaffolding (representing abstract legal concepts and relationships), provenance grounding (linking concepts to authoritative sources), and temporal navigation (maintaining temporal versions for historical queries).

The KG is a labeled directed graph G=(V,E,λ,τ) where V represents legal entities, E⊆V×V represents typed relationships, λ:V→T assigns node types, and τ:V→2^T assigns temporal validity. Node types include: *Normative* (it. “Norma”, “Principio”, “Comma/Lettera/Numero”), *Conceptual* (it. “Concetto Giuridico”, “Definizione Legale”), *Jurisprudential* (it. “Sentenza”, “Massima”, “Caso”), and *Doctrinal* (it. “Dottrina”, “Commentario”). Edge types represent: *Structural Relations* (it. “CONTIENE”, “MODIFICA”, “ABROGA”), Semantic Relations (it. “REGOLA”, “DEFINISCE”, “PRESUPPONE”), *Jurisprudential Relations* (it. “APPLICA”, “INTERPRETA”, “SOSTITUISCE”), and *Principled Relations* (it. “CONFLIGGE”, “BILANCIA”, “DERIVA\_DA”).

For temporal versioning, each Norma node *n* maintains Versions(n)={(v₁,[t₁ˢ,t₁ᵉ]),...,(vₖ,[tₖˢ,tₖᵉ])} enabling queries at specific times. The KG ensures three epistemic properties: *Explicitness* (implicit relationships become explicit edges), *Navigability* (traverse from concepts to related norms/principles/cases), and *Provenance* (every node/edge includes source metadata). While vector embeddings provide semantic similarity, explicit graph structure distinguishes relationship types, supports complex queries, and makes semantics queryable and explainable. MERL-T uses both: KG for epistemic structure, vectors for flexible retrieval.

# 4. Multi-Expert Reasoning

Legal reasoning is heterogeneous - four methodologies (literal, teleological, principled, precedential) reflect fundamentally different epistemic commitments: literal treats law as text (semiotic analysis), teleological as purposive system (intentionalist hermeneutics), principled as normative hierarchy (value reasoning), precedential as social practice (analogical reasoning). Attempting all four in a single forward pass produces epistemically confused outputs. MERL-T implements methodological specialization through four distinct experts:

1. **Literal Interpreter (legal positivism):** strict grammatical analysis and ordinary meaning (Kelsen, 1960). Prohibits considering legislative purpose, case law, or constitutional principles. Activated for clear rules, validity checks. Example: (it.) Art. 1350 c.c. requires written form for real estate - literal interpretation: oral contract is void.
2. **Systemic-teleological Reasoner (legal teleology):** identifies the *ratio legis* and interprets text to achieve purpose while maintaining systemic consistency. Activated for ambiguous text, incomplete norms. Example: (it.) Art. 1350 - why written form? Ensure certainty. Therefore interpret to include electronic signatures providing equivalent certainty.
3. **Principles Balancer (constitutional principlism):** resolves principle conflicts through proportionality analysis (legitimate aim, suitability, necessity, proportionality) (Dworkin, 1977, 1986; Alexy, 2002). Activated for constitutional questions, rights conflicts. Example: defamation balances free expression vs. honor, with journalists receiving more protection due to public interest.
4. **Precedent Analyst (legal realism):** analyzes how courts actually applied norms in analogous cases (Holmes, 1897). Extracts *ratio decidendi*, tracks temporal evolution, weights by authority. Activated for novel applications, evolving standards. Example: "good faith" in (it.) Art. 1337 c.c. requires consulting Supreme Court decisions (Bench-Capon & Sartor, 2003).

Why four experts? Each embodies distinct epistemic commitment about legal meaning - commitments mutually exclusive within a single reasoning chain. Separating experts ensures internal coherence according to epistemic commitments. Each methodology requires specialized context (e.g., only Principles Balancer receives constitutional precedents). Separation makes methodology choice visible: which expert(s) consulted, what reasoning each provided, how outputs synthesized.

# 5. LLM-Based orchestration

The Router must decide: which retrieval agents, which experts, what synthesis mode, when to iterate. Traditional MoE systems use gating networks (Jacobs et al., 1991; Shazeer et al., 2017) - non-interpretable, rigid, context-insensitive. MERL-T uses LLM-based orchestration: Router is itself an LLM reasoning about optimal execution strategy. It receives query context, enriched context, conversation history and produces *ExecutionPlan* (structured JSON) specifying *retrieval\_plan*, *reasoning\_plan*, *iteration\_strategy* with explicit rationale.

This offers four advantages: *Interpretability* (explicit rationale for every decision), Adaptability (adapts to novel queries without retraining), *Learning* (improves through feedback on execution plans), and *Compositionality* (decomposes complex queries into sub-plans). The *Iteration Controller* evaluates *ProvisionalAnswer* after each cycle based on confidence, expert consensus, norm ambiguity, and retrieval sufficiency, generating refinement plans when needed until stopping conditions are satisfied.

# 6. Synthesis: preserving epistemic plurality

The Synthesizer integrates expert outputs through two modes: *Convergent Synthesis* (when experts agree on conclusion but reason differently) extracts common conclusion, integrates multiple reasoning lines, preserves methodology attribution, and builds composite provenance. Example: minor's contract - Literal Interpreter cites Art. 2 c.c. (capacity at 18), Systemic-Teleological notes protective purpose. Synthesized answer provides both textual clarity and purposive coherence. *Divergent Synthesis* (when experts disagree substantively) identifies divergence points, explains sources, presents multiple perspectives with reasoning, and makes methodology visible. Example: journalist publishing private photos - Principles Balancer applies proportionality test (permitted if newsworthy/proportional), Precedent Analyst cites Supreme Court trend (permitted for public figures). Both favor permission with different justifications - boundary context-dependent requiring case-by-case assessment.

Why preserve disagreement? Legal ambiguity is often a genuine feature of legal materials. Forcing false consensus: (1) misrepresents epistemic state, (2) obscures methodology, (3) reduces practical value. Divergent synthesis respects epistemic plurality - tells the truth about legal complexity.

# 7. Discussion

## 7.1 Architectural necessity

We claim this architecture is necessary for trustworthy legal AI. Legal reasoning requires navigating principles/rules through multiple methodologies (Premise1). This epistemic structure cannot be adequately represented in implicit weights - it must be externalized in explicit components (Premise2). Different methodologies embody mutually exclusive commitments and must be separated for coherence (Premise3). Therefore, trustworthy legal AI must implement explicit knowledge structures (KG), separated interpretive specialists (Experts), and transparent synthesis through orchestrated pipelines. MERL-T's architecture follows necessarily from epistemic requirements of legal reasoning.

## 7.2 Comparative advantages

Compared to monolithic LLMs: explicit KG structure vs. implicit weights, four specialized methodologies vs. undifferentiated reasoning, dynamic planning vs. fixed forward pass, convergent/divergent synthesis vs. single output, full provenance vs. opacity, prompt evolution vs. retraining required, high epistemic fidelity vs. collapsed plurality. Compared to standard RAG (Lewis et al., 2020): KG enrichment guides retrieval semantically and structurally vs. pure similarity, orchestrates heterogeneous sources vs. single corpus, provides expert interpretations with provenance vs. document chunks.

## 7.3 Limitations and future work

The four-expert model captures major methodologies but not all approaches (economic analysis, critical legal studies exist). KG requires ongoing curation for new legislation/case law/doctrine (Chalkidis et al., 2020; Ashley, 2017). Multi-expert architecture requires multiple LLM calls increasing latency/cost vs. monolithic models (mitigated through parallel execution, caching, custom models for simpler steps). Current design reflects the Italian legal system - other jurisdictions require adapted architectures. Future work: cross-jurisdictional extension, production optimization, empirical validation with practitioners, integration with additional reasoning paradigms.

## 7.4 Broader implications

For Legal AI Research: architectural choices matter epistemically - system structure determines reasoning capabilities and domain fidelity. For AI Ethics (EU’s High-Level Expert Group on Artificial Intelligence, 2019): trustworthy AI requires epistemic fidelity where reasoning structure respects domain norms - MERL-T proves such fidelity can be operationalized. For Legal Practice: multi-expert architecture has pedagogical value, making methodologies explicit and comparing outputs to enhance critical thinking about legal reasoning (Allega, 2025; Allega & Puzio, 2025a).

# 8. Conclusion

Trustworthy legal AI requires architectural commitment to epistemic fidelity: system structure must mirror legal reasoning's epistemic structure. Generic LLMs cannot capture law's dual nature (principles vs. rules) or methodological plurality. MERL-T operationalizes legal epistemology through five layers: KG making implicit relationships explicit, LLM-based Router for dynamic planning, Specialized Experts grounded in distinct philosophies, Synthesis preserving epistemic plurality, and learning enabling evolution through community feedback (Allega & Puzio, 2025c). The result is a system that reasons legally - navigating between principles and rules, applying distinct methodologies, preserving uncertainty when appropriate, providing full provenance. This is not a black box but a glass box making legal reasoning transparent, auditable, and epistemically grounded. For legal AI, architecture is epistemology. The way we structure the system determines what it can know, how it can reason, and whether outputs are trustworthy. MERL-T demonstrates that architectural choices can operationalize complex domain epistemology, providing a model for trustworthy AI in specialized domains where reasoning structure matters as much as linguistic fluency.

# References

1. Alexy, R. (2002). *A theory of constitutional rights*. Oxford University Press.​
2. Allega, D. (2025). The Artificial Legal Intelligence Society as an open, multi-sided platform for law-as-computation. In M. Panait, I. G. Rădulescu, B. Tudorică, C. Popescu, & M. C. Voica (Eds.), Book of abstracts: Creativity and Innovation in Digital Economy 2025 (pp. 136–138). Petroleum-Gas University of Ploiești Publishing House. ISSN: 2971-9798
3. Allega, D., & Puzio, G. (2025a). The knowledge commoditization paradox: Theoretical and practical challenges of AI-driven value extraction in information-intensive organizations. In M. Panait, I. G. Rădulescu, B. Tudorică, C. Popescu, & M. C. Voica (Eds.), Book of abstracts: Creativity and Innovation in Digital Economy 2025 (pp. 66–68). Petroleum-Gas University of Ploiești Publishing House. ISSN: 2971-9798
4. Allega, D., & Puzio, G. (2025b). MERL-T: A multi-expert architecture for trustworthy artificial legal intelligence. In M. Panait, I. G. Rădulescu, B. Tudorică, C. Popescu, & M. C. Voica (Eds.), Book of abstracts: Creativity and Innovation in Digital Economy 2025 (pp. 170–171). Petroleum-Gas University of Ploiești Publishing House. ISSN: 2971-9798
5. Allega, D., & Puzio, G. (2025c). Reinforcement learning from community feedback (RLCF): A novel framework for artificial intelligence in social science domains. In M. Panait, I. G. Rădulescu, B. Tudorică, C. Popescu, & M. C. Voica (Eds.), Book of abstracts: Creativity and Innovation in Digital Economy 2025 (pp. 92–94). Petroleum-Gas University of Ploiești Publishing House. ISSN: 2971-9798
6. Anthropic (2024). *Claude 3 model card* (Technical Report).​ <https://assets.anthropic.com/m/61e7d27f8c8f5919/original/Claude-3-Model-Card.pdf>
7. Ashley, K. D. (2017). *Artificial intelligence and legal analytics*. Cambridge University Press.​
8. Bench-Capon, T., & Sartor, G. (2003). A model of legal reasoning with cases. *Artificial Intelligence, 150*(1-2), 97-143.​ <https://www.sciencedirect.com/science/article/pii/S0004370203001085>
9. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems 26* (pp. 2787-2795). [https://proceedings.neurips.cc/paper\_files/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf​](https://proceedings.neurips.cc/paper_files/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf%E2%80%8B)
10. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33* (pp. 1877-1901).​ <https://arxiv.org/abs/2005.14165>
11. Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: The muppets straight out of law school. In *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 2898-2904).​ <https://arxiv.org/abs/2010.02559>
12. Dworkin, R. (1977). The model of rules. In *Taking rights seriously* (pp. 14-45). Harvard University Press.​
13. Dworkin, R. (1986). *Law's empire*. Harvard University Press.​
14. Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems 30* (pp. 1024-1034).​ <https://arxiv.org/abs/1706.02216>
15. Hart, H. L. A. (1961). *The concept of law*. Oxford University Press.​
16. EU’s High-Level Expert Group on Artificial Intelligence. (2019). Ethics guidelines for trustworthy AI. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
17. Holmes, O. W. (1897). The path of the law. *Harvard Law Review, 10*(8), 457-478.​
18. Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation, 3*(1), 79-87.​ <https://www.cs.toronto.edu/~fritz/absps/jjnh91.pdf>
19. Kelsen, H. (1960). *Pure theory of law*. University of California Press.​
20. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations*.​ <https://arxiv.org/abs/1609.02907>
21. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., & Bizer, C. (2015). DBpedia – A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web Journal, 6*(2), 167-195.​ <https://www.semantic-web-journal.net/system/files/swj499.pdf>
22. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems 33* (pp. 9459-9474).​ <https://arxiv.org/abs/2005.11401>
23. Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., & Dean, J. (2017). Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *5th International Conference on Learning Representations*.​ <https://arxiv.org/abs/1701.06538>