REGULATION-AWARE LEGAL DIGITAL TWINS: CONSTRAINED WORLD MODELS FOR COUNTERFACTUAL CONTRACT PERFORMANCE, COMPLIANCE, AND DAMAGES

Anonymous authors
Paper under double-blind review

ABSTRACT

Post-deployment autonomous and agentic systems increasingly act inside socio-technical ecosystems (supply chains, trade finance networks, infrastructure projects) where factual dynamics and legal requirements are intertwined. Current LLM-centric legal tooling largely treats compliance as text generation, and therefore struggles to ground counterfactual analyses ("would this have been a breach?") or produce verifiable explanations under regulation. We propose *Regulation-Aware Legal Digital Twins* (RALDTs): constrained world models that link (i) a multi-jurisdiction legal knowledge graph to (ii) a learned commercial world model and (iii) a neuro-symbolic constraint layer used during simulation, planning, and explanation. We formalize the interface, identify a key bottleneck—mapping latent trajectories to legally salient facts—and present an implementation strategy that combines event-graph extraction with solver-guided consistency repair. Finally, we define benchmark tasks (counterfactual breach, regulation-aware planning, and explanation faithfulness) with metrics for constraint satisfaction, causal validity, uncertainty calibration, and traceability.

1 MOTIVATION

Disputes and compliance assessments are fundamentally *counterfactual*: they ask what would have happened under alternative actions, allocations, or regulatory interpretations. In cross-border contracting, liability, excuses, and damages are evaluated against a but-for baseline and quantified using legally defined remedies.(United Nations Commission on International Trade Law (UNCITRAL), 1980; UNIDROIT, 2016; CISG Advisory Council, 2008) In parallel, AI governance frameworks impose explicit operational duties (risk management, documentation, logging, human oversight) that must be satisfied *in the deployed system*, not merely described.(European Parliament and Council of the European Union, 2024; National Institute of Standards and Technology, 2023; Organisation for Economic Co-operation and Development (OECD), 2019)

Although legal language understanding has improved with domain benchmarks and specialized encoders, many high-stakes tasks still require long-horizon, dynamics-grounded reasoning.(Chalkidis et al., 2022; Koreeda & Manning, 2021; Chalkidis et al., 2020; Guha et al., 2023) Moreover, LLM outputs can be fluent yet unfaithful, with hallucination modes that are unacceptable for compliance documentation or adjudicative analysis.(Ji et al., 2022) We argue that a post-AGI compliance substrate should treat legal requirements as *constraints over trajectories* in a world model, rather than as free-form narrative generation.

2 RALDT INTERFACE AND DESIGN GOALS

Digital twins are typically defined as high-fidelity, continuously updated models that support simulation and decision-making about a real system.(Grieves, 2014; Tao et al., 2019) RALDTs adapt this paradigm to law-governed commercial ecosystems by coupling simulation with explicit normative structure and auditability.

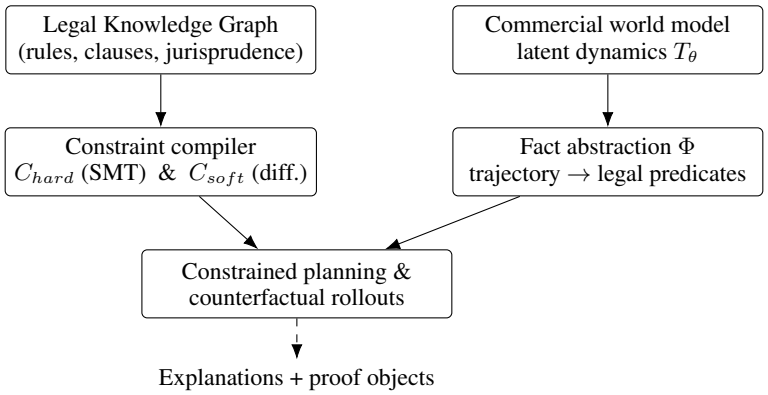

Figure 1: RALDT architecture: normative structure is compiled into constraints; a learned world model supports rollouts; a fact abstraction maps trajectories to legal predicates; planning enforces constraints and produces re-checkable explanations.

We define a RALDT as a tuple

$$\langle S, A, T_\theta, R, C, \Phi \rangle, \tag{1}$$

where S is the latent state of the commercial system, A an action space for an operator (e.g., approve shipment, deploy a model, issue notice), T_θ a learned transition model, R an operational reward (cost, delay, value), C a set of legal constraints, and Φ a *fact abstraction* mapping latent trajectories to legally salient predicates (breach, notice, causation, high-risk classification). The world model may be any latent-dynamics model that supports counterfactual rollouts and planning.(Ha & Schmidhuber, 2018; Hafner et al., 2019)

Design goals. A RALDT should support (i) *counterfactual performance*: rollouts under interventions that correspond to legally meaningful alternatives; (ii) *constraint enforcement*: reject or penalize trajectories that violate obligations; (iii) *traceability*: produce a machine-checkable record of which rules were triggered and why; and (iv) *uncertainty-aware abstention*: recognize when open-textured predicates cannot be resolved reliably.

Legal knowledge representation. We model the normative substrate as a legal knowledge graph (LKG) built on Semantic Web formalisms.(Hogan et al., 2021; Cyganiak et al., 2014; W3C OWL Working Group, 2012) Jurisdictional concepts can be grounded in legal ontologies (e.g., LKIF) to provide a shared vocabulary for cross-border reasoning.(Hoekstra et al., 2007) Graph validation and extracted-fact validation can be expressed with SHACL.(W3C RDF Data Shapes Working Group, 2017) Executable rule representations can be expressed in standards such as LegalRuleML.(Athan et al., 2015) At the logical foundation, deontic notions (obligation, permission, prohibition) motivate compiling legal rules into trajectory constraints.(von Wright, 1951)

Constraint layer. We compile the LKG into a hybrid constraint set $C = C_{\text{hard}} \cup C_{\text{soft}}$. Hard constraints cover crisp duties (deadlines, notice periods, prohibited uses) and are checked with SAT/SMT.(de Moura & Bjørner, 2008) Soft constraints cover open-textured standards (reasonableness, proportionality) and are implemented as differentiable satisfaction losses.(Serafini & d'Avila Garcez, 2016; Badreddine et al., 2022) Open texture is a design feature of legal language; thus Φ and C_{soft} must tolerate graded satisfaction and controlled abstention.(Bix, 1991)

3 Regulation as Code: From Text to Constraints

A practical bottleneck is transforming open-textured legal sources into executable constraints. "Rules as code" approaches emphasize that regulatory logic should be published in machine-consumable forms, enabling consistent implementation and audit.(Mohun & Roberts, 2020; Merigoux et al., 2021) RALDTs operationalize this idea by treating legal artifacts as *programs over trajectories*.

We propose a three-stage compilation pipeline: (1) *Canonicalization*: parse regulations, contracts, and guidance into a normalized graph schema (actors, actions, duties, exceptions) represented in RDF/OWL and validated with SHACL.(Cyganiak et al., 2014; W3C OWL Working Group, 2012; W3C RDF Data Shapes Working Group, 2017) (2) *Rule compilation*: translate rule nodes into (a) SMT constraints for crisp requirements (e.g., "retain logs for at least d days"), and (b) differentiable constraints for graded requirements (e.g., "reasonable mitigation").(de Moura & Bjørner, 2008; Badreddine et al., 2022) (3) *Linking*: connect rule predicates to facts produced by Φ so that rule evaluation is grounded in operational events.

LegalRuleML provides a principled way to represent defeasible legal rules and metadata about sources and jurisdictions.(Athan et al., 2015) For probabilistic or uncertain fact predicates, neuro-symbolic probabilistic programming can provide tractable inference while preserving rule structure.(Manhaeve et al., 2018)

4 Causality and Counterfactual Query Protocol

Counterfactual explanations should specify (i) the intervention, (ii) the causal assumptions, and (iii) the mapping from model variables to legally relevant facts. We treat counterfactual rollouts as do-interventions on a structural causal model induced from event variables and Φ.(Pearl, 2009; Peters et al., 2017; Rubin, 1974) This design supports both adjudicative counterfactuals ("but for rerouting, would breach still occur?") and normative counterfactuals ("is the decision fair under a demographic intervention?").(Kusner et al., 2017)

We propose an explicit *counterfactual query object*:

$$q = \langle \mathcal{I}, \mathcal{Y}, \mathcal{A}, \mathcal{E} \rangle,$$

where $\mathcal{I}$ is a set of interventions (actions forced or prevented), $\mathcal{Y}$ a set of target legal predicates (e.g., breach, damages triggers), $\mathcal{A}$ a set of causal assumptions (edges, independence, exchangeability), and $\mathcal{E}$ an evidence set derived from the observed trace. The twin answers by simulating T_θ under $\mathcal{I}$, abstracting facts with Φ, and checking constraints C.

5 The Bottleneck: Fact Abstraction Φ

The core technical challenge is aligning continuous, learned representations with legally salient facts. A twin can predict a delay distribution, but law asks whether a delay *constitutes breach* under a clause, whether breach *caused* a loss, or whether notice duties were triggered.(United Nations Commission on International Trade Law (UNCITRAL), 1980; UNIDROIT, 2016)

We propose a two-stage solution: (i) extract an *event graph* from operational logs and documents, using process-mining and document-structure cues to define candidate events, agents, and timestamps;(van der Aalst, 2016) (ii) learn Φ as a family of fact predictors mapping $(S_{0:T}, \text{event-graph})$ to predicate assignments, trained with *constraint-guided supervision*. During training, the solver checks C_{hard} on predicted facts and returns counterexamples that drive consistency repairs (fixing extractions or updating Φ).(de Moura & Bjørner, 2008) Soft constraints provide differentiable regularization toward legally coherent abstractions.(Serafini & d'Avila Garcez, 2016; Badreddine et al., 2022)

To prevent overconfident factual claims, we recommend calibrated uncertainty for selected predicates via conformal prediction, providing distribution-free coverage on held-out traces.(Vovk et al., 2005; Angelopoulos & Bates, 2023) This supports an operational "abstain" mode: when coverage cannot be guaranteed for a predicate under distribution shift, the system flags the predicate for human review.

6 Learning and Optimization

In practice, RALDT training couples three components: world-model learning (T_θ), fact abstraction learning (Φ_ψ), and constraint compilation/evaluation (C). Let $\tau = (S_{0:T}, A_{0:T-1})$ denote a rollout, and let $\hat{F} = \Phi_\psi(\tau)$ be predicted legal facts. We propose optimizing a constrained objective of the form

$$\max_{\theta,\psi} \mathbb{E}\Big[R(\tau) - \lambda_h \cdot \mathbb{I}\big[C_{\text{hard}}(\hat{F}) = \texttt{violate}\big] + \lambda_s \sum_{c \in C_{\text{soft}}} \texttt{sat}_c(\hat{F})\Big], \tag{2}$$

where $\texttt{sat}_c(\cdot) \in [0, 1]$ is a differentiable satisfaction score. The hard-constraint term can be implemented either as rejection sampling during planning or as a penalty with solver-generated counterexamples.(de Moura & Bjørner, 2008) Soft constraints follow Real-Logic style semantics, allowing gradients to shape Φ_ψ toward legally coherent abstractions.(Serafini & d'Avila Garcez, 2016; Badreddine et al., 2022)

Uncertainty-aware constraint checking. Many legal predicates are uncertain under incomplete evidence. For a selected subset of predicates (e.g., "high-risk deployment" classification), conformal prediction can be used to construct prediction sets with distribution-free coverage.(Vovk et al., 2005; Angelopoulos & Bates, 2023) A practical policy is to treat low-coverage predicates as "deferred" and route them to human review, while still allowing the twin to reason about downstream constraints conditionally.

From compliance to control. Regulation-aware planning can be framed as model-based control with constraints, where actions are chosen to minimize expected operational cost while respecting compiled duties.(Hafner et al., 2019; European Parliament and Council of the European Union, 2024; National Institute of Standards and Technology, 2023) This enables comparing design alternatives (contract allocation, shipping strategies, deployment configurations) under consistent regulatory semantics.

7 Benchmarks and Metrics

We propose a synthetic-but-legally-realistic benchmark suite to reduce confidentiality barriers, following dataset documentation and transparency practices.(Gebru et al., 2018; Mitchell et al., 2019) Instances are generated from transaction templates (shipping obligations, payment milestones, change orders) coupled to jurisdiction tags and parameterized clauses inspired by CISG and UNIDROIT structures.(United Nations Commission on International Trade Law (UNCITRAL), 1980; UNIDROIT, 2016) Each instance contains: (a) a clause bundle; (b) a simulated event trace; (c) an SCM over salient events; and (d) ground-truth constraint triggers.

Task 1: Counterfactual breach analysis. Given an observed trace and an intervention, determine whether breach would still occur and which duties are satisfied/violated. Score with constraint satisfaction and with *causal validity*: does the explanation correspond to a valid intervention under the SCM.(Pearl, 2009; Peters et al., 2017)

Task 2: Regulation-aware planning. Choose actions that minimize expected cost while satisfying regulatory and contractual constraints. This extends model-based control to the regulated setting.(Hafner et al., 2019; European Parliament and Council of the European Union, 2024; National Institute of Standards and Technology, 2023)

Task 3: Explanation faithfulness. Generate a natural-language explanation plus a structured proof object (predicates + evaluated constraints) that can be rechecked by a solver.(de Moura & Bjørner, 2008; Athan et al., 2015) This aligns with the motivation behind counterfactual explanation proposals in the context of automated decision-making and the GDPR.(Wachter et al., 2018; European Parliament and Council of the European Union, 2016)

8 Threat Model: Simulation Laundering

A RALDT is only as reliable as its simulator and abstractions. A motivated party can bias T_θ or Φ so that counterfactual rollouts yield legally convenient narratives—*simulation laundering*. This resembles failure modes in counterfactual explanations and recourse, where minimal-change recommendations can be brittle or non-robust.(Wachter et al., 2018; Ustun et al., 2019; Karimi et al., 2021; Dominguez-Olmedo et al., 2022)

Mitigations include: (1) requiring model cards for T_θ and Φ (intended use, jurisdictional coverage, known failure modes);(Mitchell et al., 2019) (2) requiring datasheets for benchmark templates and synthetic generation parameters;(Gebru et al., 2018) (3) adversarial stress tests over plausible

distribution shifts in event traces; and (4) separating legal reasoning from simulation assumptions by explicitly exposing SCM edges and intervention semantics.(Pearl, 2009; Peters et al., 2017)

9 Related Work (positioning)

RALDTs sit at the intersection of legal NLP benchmarks and encoders.(Chalkidis et al., 2022; 2020; Guha et al., 2023) They also draw on rules-as-code and executable-law efforts that aim to make normative logic directly implementable and auditable.(Mohun & Roberts, 2020; Merigoux et al., 2021) On the knowledge representation side, they depend on legal ontologies, rule standards, and semantic constraints for multi-jurisdiction structure.(Hoekstra et al., 2007; Athan et al., 2015; W3C RDF Data Shapes Working Group, 2017) On the modeling side, they leverage learned world models and model-based control to support counterfactual rollouts.(Ha & Schmidhuber, 2018; Hafner et al., 2019) Finally, they adopt causal and counterfactual protocols that force assumptions and interventions to be explicit.(Pearl, 2009; Peters et al., 2017; Rubin, 1974)

The counterfactual-explanation and algorithmic-recourse literatures provide operational criteria (actionability, robustness) that transfer naturally to legal counterfactuals.(Wachter et al., 2018; Ustun et al., 2019; Karimi et al., 2021; Dominguez-Olmedo et al., 2022) Our distinguishing claim is architectural: the world model, the legal graph, and the constraint layer are co-equal parts of a single decision substrate, enabling counterfactual reasoning that is grounded in dynamics and checkable against explicit normative structure.

10 Conclusion

RALDTs treat compliance as constrained control in a world model. They combine legal knowledge graphs, latent dynamics, and neuro-symbolic constraints to support counterfactual contract analysis, regulation-aware planning, and verifiable explanations. The core research agenda is robust fact abstraction: aligning learned trajectories with legally salient predicates under solver feedback and calibrated uncertainty.

Reproducibility Statement

This Tiny Paper is a proposal. We define an interface (§2), specify a compilation and causal-query protocol (§3–§4), identify the key bottleneck (Φ, §5), and propose benchmark tasks and metrics (§6). All figures are schematic; code and data are future work.

Example constraints over a trajectory
Hard: `notice_given` $\rightarrow t_{\text{notice}} \leq t_{\text{deadline}}$
Hard: `high_risk_deployed` $\rightarrow$ `logs_retained` $\wedge$ `human_oversight`
Soft: `reasonable_mitigation` $= \sigma(w^\top f(S_{0:T}))$
Causal: `loss` $\leftarrow do($`reroute` $= 1)$ is evaluated under SCM

Figure 2: Constraint types: crisp duties checked by SMT, graded standards with differentiable satisfaction, and causal queries requiring explicit interventions.

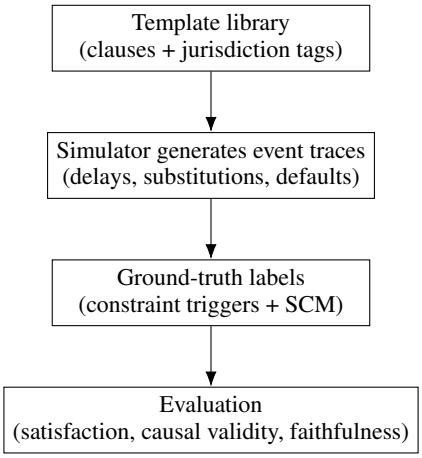

Figure 3: Benchmark generation pipeline for reproducible RALDT evaluation, using synthetic templates to avoid confidential data.

A Additional figures (schematic)

References

Anastasios N. Angelopoulos and Stephen Bates. Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4):494–591, 2023. doi: 10.1561/2200000101.

Tara Athan, Guido Governatori, Monica Palmirani, Adrian Paschke, and Adam Wyner. Legalruleml: Design principles and foundations. In *Reasoning Web. Web Logic Rules*, volume 9203 of *Lecture Notes in Computer Science*, pp. 151–188. Springer, 2015. doi: 10.1007/978-3-319-21768-0_6.

Samy Badreddine, Artur d'Avila Garcez, Luciano Serafini, and Michael Spranger. Logic tensor networks. *Artificial Intelligence*, 303:103649, 2022. doi: 10.1016/j.artint.2021.103649.

Brian Bix. H. l. a. hart and the "open texture" of language. *Law and Philosophy*, 10:51–72, 1991. doi: 10.1007/BF00144295.

Ilias Chalkidis, Manos Fergadiotis, Prodromos Malakasiotis, Nikolaos Aletras, and Ion Androutsopoulos. LEGAL-BERT: The muppets straight out of law school. In Trevor Cohn, Yulan He, and Yang Liu (eds.), *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 2898–2904, Online, November 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.findings-emnlp.261. URL https://aclanthology.org/2020.findings-emnlp.261/.

Ilias Chalkidis, Abhik Jana, Dirk Hartung, Michael Bommarito, Ion Androutsopoulos, Daniel Katz, and Nikolaos Aletras. LexGLUE: A benchmark dataset for legal language understanding in English. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio (eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 4310–4330, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-long.297. URL https://aclanthology.org/2022.acl-long.297/.

CISG Advisory Council. Cisg advisory council opinion no. 8: Calculation of damages under cisg articles 75 and 76, 2008. URL https://cisgac.com/opinion-no8/. Adopted 15 November 2008.

Richard Cyganiak, David Wood, and Markus Lanthaler. Rdf 1.1 concepts and abstract syntax. W3C Recommendation, 2014. URL https://www.w3.org/TR/rdf11-concepts/.

Leonardo de Moura and Nikolaj Bjørner. Z3: An efficient smt solver. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2008)*, volume 4963 of *Lecture Notes in Computer Science*, pp. 337–340. Springer, 2008. doi: 10.1007/978-3-540-78800-3_24.

Ricardo Dominguez-Olmedo, Amir-Hossein Karimi, and Bernhard Schölkopf. On the adversarial robustness of causal algorithmic recourse. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 5324–5342, 2022. URL https://proceedings.mlr.press/v162/dominguez-olmedo22a.html.

European Parliament and Council of the European Union. Regulation (eu) 2016/679 (general data protection regulation), 2016. URL https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. OJ L 119, 4 May 2016.

European Parliament and Council of the European Union. Regulation (eu) 2024/1689 laying down harmonised rules on artificial intelligence (artificial intelligence act), 2024. URL https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng. OJ L, 2024/1689, 12 July 2024.

Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. Datasheets for datasets. arXiv preprint arXiv:1803.09010, 2018. URL https://arxiv.org/abs/1803.09010.

Michael Grieves. Digital twin: Manufacturing excellence through virtual factory replication. White paper, 2014. URL https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication.

Neel Guha, Julian Nyarko, Daniel E. Ho, Adam Chilton, Arvind Narayanan, Alex Pentland, et al. Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models. arXiv preprint arXiv:2308.11462, 2023. URL https://arxiv.org/abs/2308.11462.

David Ha and Jürgen Schmidhuber. World models. arXiv preprint arXiv:1803.10122, 2018. URL https://arxiv.org/abs/1803.10122.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. arXiv preprint arXiv:1912.01603, 2019. URL https://arxiv.org/abs/1912.01603.

Rinke Hoekstra, Joost Breuker, Marcello Di Bello, and Alexander Boer. The lkif core ontology of basic legal concepts. In *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, volume 321 of *CEUR Workshop Proceedings*, pp. 43–63, 2007. URL https://ceur-ws.org/Vol-321/paper3.pdf.

Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, Sabrina Kirrane, Jose Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, Axel-Cyrille Ngonga Ngomo, Axel Polleres, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, and Antoine Zimmermann. Knowledge graphs. *ACM Computing Surveys*, 54(4):71:1–71:37, 2021. doi: 10.1145/3447772.

Ziwei Ji, Nayeon Lee, Rita Frieske, Tiancheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Yuan Bang, Andrea Madotto, and Pascale Fung. A survey of hallucination in natural language generation. arXiv preprint arXiv:2202.03629, 2022. URL https://arxiv.org/abs/2202.03629.

Amir-Hossein Karimi, Bernhard Schölkopf, and Isabel Valera. Algorithmic recourse: from counterfactual explanations to interventions. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2021. doi: 10.1145/3442188.3445899.

Yuta Koreeda and Christopher Manning. ContractNLI: A dataset for document-level natural language inference for contracts. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih (eds.), *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 1907–1919, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.findings-emnlp.164. URL https://aclanthology.org/2021.findings-emnlp.164/.

Matt J. Kusner, Joshua R. Loftus, Chris Russell, and Ricardo Silva. Counterfactual fairness. In *Advances in Neural Information Processing Systems*, 2017. URL https://papers.nips.cc/paper/6995-counterfactual-fairness.

Robin Manhaeve, Sebastijan Dumančić, Angelika Kimmig, Thomas Demeester, and Luc De Raedt. Deepproblog: Neural probabilistic logic programming. In *Advances in Neural Information Processing Systems*, 2018. URL https://papers.nips.cc/paper/7750-deepproblog-neural-probabilistic-logic-programming.

Denis Merigoux, Nicolas Chataing, and Jonathan Protzenko. Catala: A programming language for the law. In *Proceedings of the ACM on Programming Languages (ICFP 2021)*, 2021. doi: 10.1145/3473582. URL https://icfp21.sigplan.org/details/icfp-2021-papers/16/Catala-A-Programming-Language-for-the-Law.

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting. In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2019. doi: 10.1145/3287560.3287596. URL https://arxiv.org/abs/1810.03993.

James Mohun and Alex Roberts. Cracking the code: Rulemaking for humans and machines. Technical Report OECD Working Papers on Public Governance, No. 42, OECD, 2020. URL https://www.oecd-ilibrary.org/governance/cracking-the-code-rulemaking-for-humans-and-machines_3afe6ba5-en.

National Institute of Standards and Technology. Artificial intelligence risk management framework (ai rmf 1.0). Technical Report NIST AI 100-1, U.S. Department of Commerce, 2023. URL https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.

Organisation for Economic Co-operation and Development (OECD). Recommendation of the council on artificial intelligence, 2019. URL https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2 edition, 2009.

Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT Press, 2017. URL https://web.math.ku.dk/~peters/jonas_files/ElementsOfCausalInference.pdf.

Donald B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies, 1974.

Luciano Serafini and Artur d'Avila Garcez. Logic tensor networks: Deep learning and logical reasoning from data and knowledge. arXiv preprint arXiv:1606.04422, 2016. URL https://arxiv.org/abs/1606.04422.

Fei Tao, Qinglin Qi, Lihui Wang, and A. Y. C. Nee. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering*, 5(4):653–661, 2019. doi: 10.1016/j.eng.2019.01.014.

UNIDROIT. Unidroit principles of international commercial contracts 2016, 2016. URL https://www.unidroit.org/wp-content/uploads/2021/06/Unidroit-Principles-2016-English-bl.pdf.

United Nations Commission on International Trade Law (UNCITRAL). United nations convention on contracts for the international sale of goods (cisg), 1980. URL https://uncitral.un.org/en/texts/salegoods/conventions/sale_of_goods/cisg. Adopted 11 April 1980.

Berk Ustun, Alexander Spangher, and Yang Liu. Actionable recourse in linear classification. In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2019. doi: 10.1145/3287560.3287566. URL https://dl.acm.org/doi/10.1145/3287560.3287566.

Wil M. P. van der Aalst. *Process Mining: Data Science in Action*. Springer, 2 edition, 2016. doi: 10.1007/978-3-662-49851-4.

Georg Henrik von Wright. Deontic logic. *Mind*, 60(237):1–15, 1951. doi: 10.1093/mind/LX.237.1.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic Learning in a Random World*. Springer, 2005. doi: 10.1007/b106715.

W3C OWL Working Group. Owl 2 web ontology language document overview (second edition). W3C Recommendation, 2012. URL https://www.w3.org/TR/owl2-overview/.

W3C RDF Data Shapes Working Group. Shapes constraint language (shacl). W3C Recommendation, 2017. URL https://www.w3.org/TR/shacl/.

Sandra Wachter, Brent Mittelstadt, and Chris Russell. Counterfactual explanations without opening the black box: Automated decisions and the gdpr. *Harvard Journal of Law & Technology*, 31(2):841–887, 2018. URL https://jolt.law.harvard.edu/assets/articlePDFs/v31/Counterfactual-Explanations-without-Opening-the-Black-Box-Sandra-Wachter-et-al.pdf.