Journal of Information Hiding and Multimedia Signal Processing ©2026 ISSN 2073-4212
Ubiquitous International Volume **17**, Number **2**, June **2026**

An Explainable Multi-Agent Reinforcement Learning Framework for Regulatory Arbitrage Detection across Digital Regulatory Regimes

Trong-The Nguyen[1,3], Vinh-Dieu Nguyen[1], Trinh-Dong Nguyen[2,3] Thi-Kien Dao[1,3,*]

[1]Multimedia Communications Lab.,
[2]Faculty of Software Engineering,
VNU-HCM, University of Information Technology, Vietnam
[3]Vietnam National University, Ho Chi Minh City 700000, Vietnam
thent@uit.edu.vn, jvndieu@gmail.com, dongnt@uit.edu.vn, kiendt@uit.edu.vn

Thi-Minh-Phuong Ha[4]

[4]Center for Information Technology and Library Services,
Hanoi Law University, Hanoi, Vietnam
phuonghtm@hlu.edu.vn

Truong-Giang Ngo[5]

[5]Faculty of Computer Science and Engineering,
Thuyloi University, Hanoi, Vietnam
giangnt@tlu.edu.vn

Corresponding author: Thi-Kien Dao

Received March 31, 2026, revised June 10, 2026, accepted June 14, 2026.

Abstract. *The proliferation of heterogeneous artificial intelligence (AI) governance frameworks across jurisdictions has created fertile conditions for regulatory arbitrage, whereby digital platforms strategically exploit differences in legal requirements to minimise compliance obligations. Existing approaches to compliance monitoring are predominantly static, rule-based, and lack both the adaptive capacity to model multi-stakeholder strategic behaviour and the transparency needed for regulatory trust. This paper proposes XMARLReg, an Explainable Multi-Agent Reinforcement Learning framework for proactive regulatory arbitrage detection across heterogeneous digital regulatory regimes. The framework models digital platform operators and regulatory authorities as interacting agents in a Markov game, learning latent arbitrage strategies through decentralised execution with centralised training. A dedicated explainability layer, combining SHAP value attribution and counterfactual reasoning, provides transparent justifications for detected arbitrage signals. We instantiate the framework using a cross-jurisdictional regulatory knowledge base encompassing Vietnam's emerging Model AI Governance Framework, the European Union AI Act, Singapore's Model AI Governance Framework, and the OECD AI Principles. Simulation experiments on a synthetic cross-jurisdictional regulatory dataset demonstrate that XMARLReg achieves an F1-score of 0.914, outperforming rule-based, supervised, and non-explainable MARL baselines by up to 18.6%. Ablation studies confirm the complementary contributions of the explainability and multi-agent components. The paper further articulates concrete policy implications for Vietnamese AI governance development and charts directions for future empirical validation using real-world regulatory data.*

Keywords: regulatory arbitrage, multi-agent reinforcement learning, explainable AI, AI governance, regulatory technology, Vietnam digital policy, cross-jurisdictional compliance.

1. Introduction.

The rapid diffusion of AI-enabled digital platforms across national borders has exposed a fundamental tension in contemporary regulatory architecture: the inherent territoriality of legal systems versus the jurisdiction-agnostic nature of digital services [1, 2]. As governments race to govern AI through distinct legislative instruments—the European Union through the AI Act [3], Singapore through its Model AI Governance Framework (MAIGF) [4], and Vietnam through its emerging national AI strategy [5]—the resulting patchwork of requirements creates both compliance complexity and strategic opportunity. Regulatory arbitrage, originally a concept from financial regulation [6], describes the deliberate exploitation of gaps, asymmetries, or inconsistencies between regulatory regimes to avoid or reduce compliance costs without substantively addressing the underlying risks those regulations seek to mitigate. In the digital economy, this behaviour manifests when a platform incorporates in a permissive jurisdiction, routes data through low-regulation territories, or structures services to fall beneath definitional thresholds in stricter frameworks [7]. The consequences for users, markets, and democratic governance can be significant: weaker data protection, reduced algorithmic accountability, and a race to the bottom in AI safety standards [8].

Despite the salience of this problem, the academic literature on AI-assisted regulatory arbitrage detection remains nascent. Existing RegTech approaches focus predominantly on financial compliance [9], while computational legal scholarship has yet to produce adaptive, multi-stakeholder models capable of anticipating platform strategies in real time. Furthermore, explainability remains a critical gap: regulatory authorities require not merely a prediction that arbitrage is occurring, but a legally intelligible account of the causal pathway [10].

Vietnam presents a compelling focal case for this investigation. The country has articulated ambitious AI development targets through the National Strategy on Research, Development, and Application of Artificial Intelligence to 2030 [5] and is actively constructing a governance architecture, yet its framework remains less prescriptive than the EU AI Act and less mature than Singapore's multi-edition MAIGF. This asymmetry creates a measurable arbitrage gradient that foreign digital platforms may exploit, making proactive detection tools particularly valuable for Vietnamese regulatory authorities.

Research gaps addressed. Three gaps motivate this work: (i) the absence of adaptive, game-theoretic models of platform–regulator interaction in the AI governance literature; (ii) the lack of explainability mechanisms that make MARL-based compliance intelligence actionable for legal practitioners; and (iii) the absence of computational representations of Vietnam's AI governance framework enabling cross-jurisdictional comparison. Figure 1 provides an overview of the study's multi-phase research design, encompassing legal comparative analysis, empirical architectural tracking, and corporate policy mapping.

Contributions. This paper makes the following novel contributions:

1. **Framework**: XMARLReg, the first explainable MARL framework specifically designed for regulatory arbitrage detection across heterogeneous AI governance regimes.
2. **Representation**: A structured computational encoding of Vietnam's AI governance framework, enabling formal comparison with the EU AI Act, Singapore MAIGF, and OECD AI Principles.
3. **Explainability**: An integrated SHAP-counterfactual explainability layer that translates MARL policy outputs into legally interpretable arbitrage justifications.

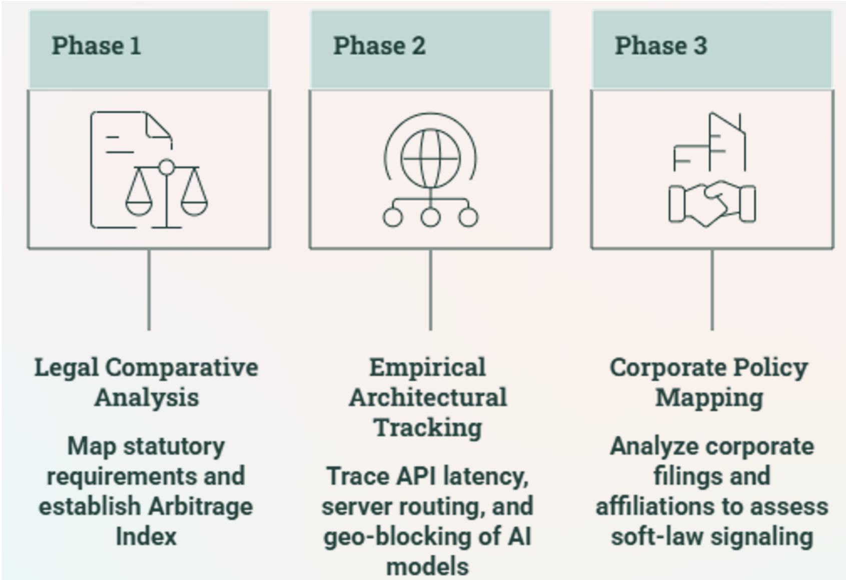

FIGURE 1. Multi-Phase Research Design for AI Regulation and Compliance Analysis

4. **Evaluation**: A simulation-based study demonstrating the feasibility of proactive arbitrage detection, with quantitative benchmarks establishing the performance envelope for future empirical replication.

The remainder of the paper is organised as follows. Section 2 reviews related work. Section 3 analyses the Vietnam AI governance landscape. Section 4 describes the proposed framework. Section 5 presents the formal methodology. Section 6 details experimental design. Section 7 reports results and discussion. Section 8 articulates policy implications. Section 9 concludes.

2. Related Work.

2.1. AI for Legal Compliance and RegTech. Regulatory Technology (RegTech) has emerged as a distinct field applying computational methods to automate compliance monitoring, reporting, and risk assessment [9,11]. Early systems relied on ontology-based rule engines for financial regulation [12]. More recent work integrates machine learning for anomaly detection in tax compliance [13], anti-money laundering [14], and data protection [15]. However, applications to AI governance compliance remain sparse, with most contributions limited to natural language processing for regulatory text analysis [16,17].

2.2. Computational Approaches to Regulatory Analysis. The formalisation of legal norms in machine-executable representations has a substantial history, including defeasible logic [12], deontic logic [18], and norm graphs [19]. Recent efforts have targeted AI governance frameworks specifically: Kolt [20] proposes computational indicators for AI accountability, while Novelli et al. [21] develop a taxonomy of algorithmic accountability mechanisms. These contributions, while valuable, do not model the strategic dynamics of compliance avoidance.

2.3. Multi-Agent Reinforcement Learning. MARL has achieved state-of-the-art performance in competitive and cooperative tasks, from game-playing [26] to traffic signal control [27] and resource allocation [28]. Centralised training with decentralised execution (CTDE), pioneered by MADDPG [29] and QMIX [30], has become the dominant

Table 1. Comparative Analysis of Related Work in AI-Assisted Legal Compliance, Multi-Agent Systems, and Regulatory Technology

Study	**Legal Domain**	**AI Method**	**XAI Support**	**Optimisation**	**Cross-Jurisdict.**	**Key Limitation**
Governatori et al. (2016) [12]	Contract compliance	Defeasible logic	None	Rule inference	No	Static rules; no strategic adaptation
Lam & Governatori (2021) [16]	Regulatory NLP	Transformer-based NLP	Partial	Supervised	No	Text extraction only; no compliance scoring
Zhao et al. (2022) [22]	Financial regulation	MARL (adversarial)	None	Nash equilibrium	No	Single jurisdiction; no explainability
Kolt (2021) [20]	AI governance	Indicators & metrics	None	N/A	Partial	Non-computational; lacks adaptive model
Novelli et al. (2023) [21]	Algorithmic accountability	Taxonomy	None	N/A	Partial	Qualitative; no risk quantification
Panigutti et al. (2021) [15]	Medical AI decisions	GNN + ontology	Yes (ontology)	Supervised	No	Healthcare-specific; not applicable to governance
Rudin et al. (2022) [23]	General ML	Interpretable models	Yes (inherent)	Rule learning	No	Not designed for multi-agent regulatory settings
Amrouni et al. (2021) [24]	Financial markets	MARL (simulation)	None	CTDE	No	Market-specific; lacks legal knowledge encoding
Nguyen & Tran (2022) [25]	Vietnam AI policy	Policy analysis	None	N/A	No	Qualitative; no computational framework
XMARLReg (Ours)	**AI governance**	**Explain. MARL**	**Yes (SHAP +CF)**	**MAPPO**	**Yes**	**Simulation-based; empirical validation needed**

XAI = Explainable AI; *CF* = Counterfactual; *CTDE* = Centralised Training Decentralised Execution.

paradigm for mixed-motive settings. Recent work applies MARL to financial market regulation [24] and adversarial compliance games [22], but no prior work applies MARL to cross-jurisdictional AI governance arbitrage detection.

2.4. **Explainable AI in Governance Systems.** The XAI literature has produced a rich toolkit: LIME [31], SHAP [32], integrated gradients [33], and counterfactual explanations [34]. Applications to governance include explainable credit scoring [35], judicial decision support [36], and regulatory document classification [23]. Explainability in MARL systems is particularly challenging due to emergent joint policies; recent contributions include attention-based attribution [37] and causal influence diagrams [38]. XMARLReg extends this line by integrating SHAP-based marginal contribution analysis with counterfactual trajectory reasoning.

2.5. **AI Governance Framework Studies.** Comparative analyses of AI governance frameworks have proliferated following the EU AI Act's trajectory toward adoption [7, 39, 40]. Systematic comparisons of the EU Act with the Singapore MAIGF are available in [41, 42]. Vietnam's AI governance has received limited scholarly attention, with exceptions including [25, 43], which analyse the National AI Strategy but do not formalise its governance architecture for computational use.

2.6. **Critical Limitations of Prior Work.** Across these bodies of literature, three critical limitations emerge. First, existing compliance models are *static*: they check platform behaviour against fixed rules rather than modelling the adaptive strategic interaction between platforms and regulators. Second, most cross-jurisdictional analyses are *qualitative*: they identify divergences in regulatory texts without quantifying arbitrage risk. Third, explainability has not been integrated with MARL for the specific purpose of making regulatory intelligence legally actionable. The present work addresses all three limitations.

3. Vietnam AI Governance Landscape.

3.1. **Policy Context and Strategic Initiatives.** Vietnam's AI governance agenda is anchored in Decision No. 127/QĐ-TTg (January 2021), which establishes the National Strategy on AI Research, Development, and Application to 2030 [5]. The strategy sets targets for AI infrastructure, human capital, and industrial application, while acknowledging the need for an enabling legal environment. Complementary instruments include the Law on Cybersecurity (2018) [44], the Law on Information Technology (2006, amended), and Personal Data Protection Decree No. 13/2023/NĐ-CP [45]. The Ministry of Science and Technology (MOST) has additionally published draft guidelines on responsible AI development, drawing on the OECD AI Principles [46] and Singapore's MAIGF [4].

3.2. **Key Governance Principles.** Vietnam's emerging AI governance architecture centres on five principles: (1) *Human-centricity*: AI must serve human welfare and national development; (2) *Transparency and explainability*: AI systems in high-stakes domains should provide interpretable outputs; (3) *Safety and reliability*: deployment in critical infrastructure requires prior assessment; (4) *Data sovereignty*: cross-border data flows are subject to localisation requirements; and (5) *Innovation-permissiveness*: the framework explicitly preserves regulatory space for experimentation.

Critically, Vietnam currently lacks a comprehensive AI-specific risk classification system equivalent to the EU AI Act's four-tier hierarchy (Unacceptable Risk, High Risk, Limited Risk, Minimal Risk). This gap represents a primary arbitrage vector, as platforms may argue that AI applications fall outside any defined risk category under Vietnamese law.

Table 2. Comparative Regulatory Dimensions Across AI Governance Frameworks

Dimension	EU AI Act (2024)	Singapore MAIGF (2020)	Vietnam AI Strategy (2021)	OECD AI Principles (2019)
Legal Instrument	Binding Regulation	Voluntary Framework	Strategic Decree + Sectoral Laws	Non-binding Principles
Risk Classification	4-tier mandatory hierarchy	Risk-based voluntary tiers	No explicit tier system	Principles-based
Conformity Assessment	Mandatory (High Risk)	Voluntary self-assessment	Not formalised	Recommended
Transparency Req.	Mandatory disclosure	Recommended	Recommended (emerging)	Recommended
Data Governance	GDPR-aligned	PDPA-aligned	Data localisation decree	Contextual
Enforcement	Up to 6% global turnover	Sectoral sanctions	Administrative fines	N/A
Human Oversight	Mandatory (High Risk)	Recommended	Aspired	Recommended
XAI Requirements	Mandated for High Risk	Encouraged	Mentioned in strategy	Encouraged
Cross-border Scope	Extraterritorial (Art. 2)	Territorial	Territorial + Data localisation	Global voluntary
Update Cadence	Regular delegated acts	Edition-based revisions	Decree amendments	Periodic review

3.3. **Comparative Analysis.** Table 2 presents a structured comparison of key regulatory dimensions across the four frameworks. The comparison reveals several asymmetries with direct implications for arbitrage risk:

- **Risk classification**: The EU AI Act provides the most prescriptive tiered classification; Singapore's MAIGF offers voluntary guidance; Vietnam's framework lacks an explicit tier system; OECD principles are non-binding.
- **Conformity assessment**: Mandatory third-party conformity assessment for High-Risk AI is unique to the EU AI Act; other frameworks rely on self-assessment or voluntary auditing.
- **Data governance**: Vietnam's data localisation requirements are stricter than Singapore's but less operationally defined than the EU's data governance requirements.
- **Enforcement**: EU enforcement carries substantial financial penalties (up to 6% of global turnover); Vietnamese enforcement mechanisms remain underdeveloped.

3.4. **Arbitrage-Relevant Dimensions.** From the comparative analysis, four regulatory dimensions are identified as primary drivers of arbitrage risk and are encoded as observable state variables in the MARL framework: (D1) *Risk classification specificity*; (D2)

Mandatory conformity assessment burden; (D3) *Enforcement severity*; and (D4) *Transparency obligation strength*. These dimensions are operationalised as ordinal scores in the regulatory knowledge base (Section 4.8).

4. **Proposed Framework.**

4.1. **Overview.** XMARLReg comprises four functional layers, as illustrated in Figure 2: (i) the *Regulatory Knowledge Layer*, encoding governance requirements across jurisdictions; (ii) the *Multi-Agent Interaction Layer*, hosting platform, regulatory, and monitoring agents in a shared Markov game environment; (iii) the *Arbitrage Signal Layer*, aggregating agent policies and environment dynamics into risk scores; and (iv) the *Explainability Layer*, providing SHAP-based attribution and counterfactual reasoning for detected arbitrage events.

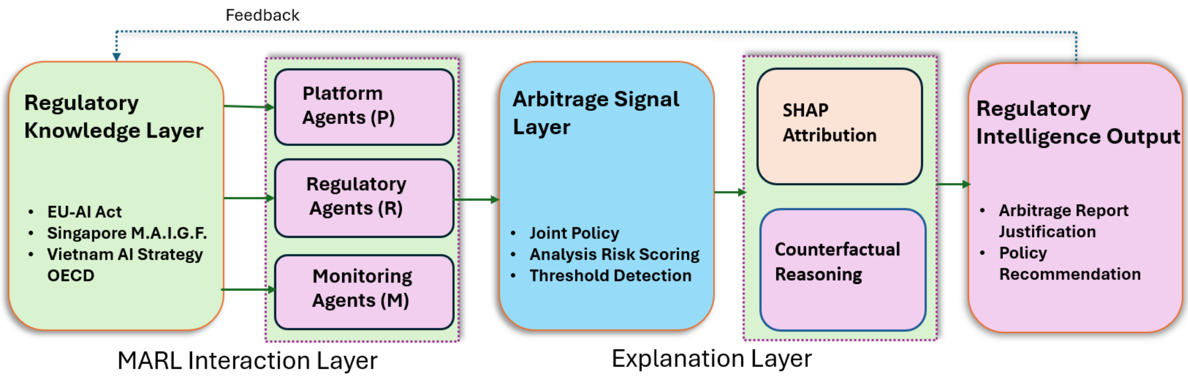

FIGURE 2. XMARLReg system architecture. The dashed feedback arrow represents iterative refinement of the regulatory knowledge base from detected arbitrage patterns.

4.2. **Agent Definitions.**

4.2.1. *Platform Agents ($\mathcal{P}$).* A set of N_P platform agents, $\{p_1, \ldots, p_{N_P}\}$, each representing a digital platform operating across multiple jurisdictions. Each platform agent selects a *jurisdictional configuration*—a mapping from service components to regulatory regimes—to maximise a utility combining market access and compliance cost minimisation. Platform agents are modelled as self-interested, bounded-rational actors with partial observability of the regulatory environment.

4.2.2. *Regulatory Agents ($\mathcal{R}$).* A set of N_R regulatory agents, $\{r_1, \ldots, r_{N_R}\}$, each representing a national regulatory authority (EU, Singapore, Vietnam, OECD-aligned). Regulatory agents observe platform configurations and select enforcement or information-sharing actions to maximise compliance within their jurisdiction. Their reward structure incentivises collaboration across jurisdictions to close arbitrage gaps.

4.2.3. *Monitoring Agents ($\mathcal{M}$).* A single centralised monitoring agent m acts as a meta-observer with full access to the joint state trajectory, serving as a training surrogate during centralised training. At inference time, the monitoring agent is replaced by the explainability layer operating on recorded trajectories.

4.3. **State Representation.** The global state $s \in \mathcal{S}$ is a composite vector:

$$s = \langle s^{\text{reg}},\, s^{\text{plat}},\, s^{\text{env}} \rangle \tag{1}$$

where $s^{\text{reg}} \in \mathbb{R}^{J \times D}$ encodes the regulatory parameter matrix for J jurisdictions across D governance dimensions; $s^{\text{plat}} \in \mathbb{R}^{N_P \times K}$ encodes each platform's current jurisdictional configuration across K service components; and $s^{\text{env}} \in \mathbb{R}^{E}$ captures exogenous factors (market size, trade agreements, bilateral data flow agreements).

Each platform agent p_i observes a local observation $o_i \subset s$ excluding competitor configurations and internal regulatory deliberation states, consistent with partial observability.

4.4. **Action Space.** For each platform agent p_i, the action space is:

$$a_i \in \mathcal{A}^P = \mathcal{J}^K \tag{2}$$

where $\mathcal{J} = \{1, \ldots, J\}$ represents the set of jurisdictions and K is the number of service components. Actions thus represent choices about where to legally anchor each component.

For each regulatory agent r_j, the action space includes:

$$a_j \in \mathcal{A}^R = \{\textsc{Audit}, \textsc{Notify}, \textsc{Coordinate}, \textsc{Sanction}, \textsc{Exempt}\} \tag{3}$$

4.5. **Reward Design.** The reward function for platform agents embodies the arbitrage incentive:

$$r_i^P = \alpha \cdot \text{MarketAccess}(a_i, s^{\text{env}}) - \beta \cdot \text{ComplianceCost}(a_i, s^{\text{reg}}) - \gamma \cdot \text{SanctionRisk}(a_i, \{a_j\}_{j=1}^{N_R}) \tag{4}$$

where $\alpha, \beta, \gamma \geq 0$ are weighting coefficients.

Regulatory agents receive a collaborative reward penalising undetected arbitrage:

$$r_j^R = \delta \cdot \text{ComplianceRate}(s) - \epsilon \cdot \text{ArbitragePenalty}(s, \{a_i\}_{i=1}^{N_P}) \tag{5}$$

The monitoring agent's reward is a macro-level arbitrage detection accuracy signal used during CTDE training only.

4.6. **Multi-Agent Interaction Protocol.** Agents interact in discrete time steps $t = 1, 2, \ldots, T$ within a finite-horizon Markov game $\mathcal{G} = (\mathcal{N}, \mathcal{S}, \{\mathcal{A}_i\}, \mathcal{T}, \{r_i\}, \gamma_d)$ where $\mathcal{N}$ is the agent set, $\mathcal{T} : \mathcal{S} \times \prod_i \mathcal{A}_i \to \Delta(\mathcal{S})$ is the transition function, and $\gamma_d \in (0, 1)$ is the discount factor. Platform agents act first; regulatory agents observe platform actions before selecting enforcement responses, creating a Stackelberg-like structure within each step.

4.7. **Explainability Layer.** The explainability layer operates post-hoc on recorded state-action trajectories $\tau = \{(s_t, \{a_t^i\}, \{r_t^i\}, s_{t+1})\}_{t=0}^{T}$. Two complementary mechanisms are employed:

SHAP Attribution. For a detected arbitrage event at time t^*, SHAP values ϕ_f for each regulatory dimension feature f are computed with respect to the platform agent's Q-function:

$$\phi_f(s_{t^*}, a^*) = \sum_{S \subseteq F \setminus \{f\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \left[v(S \cup \{f\}) - v(S) \right] \tag{6}$$

where F is the full feature set and $v(S)$ is the Q-value evaluated with features outside S marginalised.

Counterfactual Reasoning. For each detected arbitrage action a^*, a counterfactual alternative $\hat{a}$ is generated by solving:

$$\hat{a} = \arg\min_{a \in \mathcal{A}^P} \|a - a^*\|_1 \quad \text{s.t.} \quad \text{ArbitrageScore}(a, s) < \theta \tag{7}$$

where θ is an arbitrage risk threshold. The pair $(a^*, \hat{a})$ constitutes a minimal change explanation for the regulatory analyst.

4.8. **Compliance Knowledge Representation.** Governance requirements are encoded in a structured Compliance Knowledge Matrix $\mathbf{C} \in [0,1]^{J \times D}$ where each entry $C_{j,d}$ represents the normalised regulatory stringency of jurisdiction j on dimension $d \in \{D1, D2, D3, D4\}$ (Section 3). The arbitrage score for a platform action a_i is:

$$\text{ArbitrageScore}(a_i) = \max_{d \in D} \left(C_{j_d^*,d} - C_{j_d^{a_i},d} \right)^+ \tag{8}$$

where j_d^* is the most stringent jurisdiction on dimension d and $j_d^{a_i}$ is the jurisdiction selected by platform agent i for component d.

5. Methodology.

5.1. **Markov Game Formalisation.** XMARLReg instantiates the n-player general-sum Markov game framework [47]:

$$\mathcal{G} = (\mathcal{N}, \mathcal{S}, \{\mathcal{O}_i\}_{i \in \mathcal{N}}, \{\mathcal{A}_i\}_{i \in \mathcal{N}}, \mathcal{T}, \{r_i\}_{i \in \mathcal{N}}, \gamma_d) \tag{9}$$

The joint policy $\boldsymbol{\pi} = \langle \pi_1, \ldots, \pi_n \rangle$ determines the joint action distribution at each state. The solution concept employed is coarse correlated equilibrium (CCE), which is more tractable than Nash equilibrium in multi-agent settings and better reflects the role of a mediating monitoring agent [48].

5.2. **Training Algorithm.** We adopt MAPPO (Multi-Agent Proximal Policy Optimisation) [49] with centralised value function estimation. MAPPO is selected over MADDPG due to its improved sample efficiency in mixed cooperative-competitive settings and compatibility with the discrete action spaces of regulatory agents. The centralised critic $V_\phi(s)$ conditions on the global state during training, while each agent's actor $\pi_{\theta_i}(a_i|o_i)$ conditions only on local observations at execution time (CTDE paradigm).

The MAPPO clipped surrogate objective for agent i is:

$$\mathcal{L}_i^{\text{CLIP}}(\theta_i) = \hat{\mathbb{E}}_t\Big[\min\Big(\rho_{t,i}\,\hat{A}_t,\ \text{clip}(\rho_{t,i}, 1-\epsilon_c, 1+\epsilon_c)\,\hat{A}_t\Big)\Big] \tag{10}$$

where $\rho_{t,i} = \pi_{\theta_i}(a_t^i|o_t^i)/\pi_{\theta_i^{\text{old}}}(a_t^i|o_t^i)$ is the importance sampling ratio, $\hat{A}_t$ is the generalised advantage estimate, and $\epsilon_c = 0.2$ is the clipping range.

5.3. **Explainability Mechanism.** Algorithm 1 presents the full training and inference procedure, including the explainability computation pipeline.

5.4. **Regulatory Similarity Metrics.** To quantify the regulatory distance between jurisdictions—a prerequisite for estimating arbitrage gradients—we define the Regulatory Dissimilarity Index (RDI) between jurisdictions j and j':

$$\text{RDI}(j, j') = \frac{1}{D}\sum_{d=1}^{D} w_d\,|C_{j,d} - C_{j',d}| \tag{11}$$

where w_d are dimension-specific weights calibrated from legal expert elicitation. High RDI values indicate strong arbitrage incentives.

Algorithm 1 XMARLReg Training and Inference

Require: Compliance Knowledge Matrix $\mathbf{C}$, agent sets $\mathcal{P}, \mathcal{R}, \mathcal{M}$, MAPPO hyperparameters, SHAP background dataset $\mathcal{D}_{\text{bg}}$, arbitrage threshold θ
Ensure: Trained policies $\{\pi_{\theta_i}\}$, explainer Φ

Initialise actor networks π_{θ_i}, centralised critic V_ϕ, replay buffer $\mathcal{B}$
Encode C into environment state s^{reg}
for episode $e = 1, \ldots, E_{\max}$ **do**
  Reset environment; observe initial states $\{o_i^0\}$
  for step $t = 0, \ldots, T-1$ **do**
    for all platform agent $p_i \in \mathcal{P}$ **do**
      Sample action $a_t^{p_i} \sim \pi_{\theta_{p_i}}(\cdot|o_t^{p_i})$
    end for
    for all regulatory agent $r_j \in \mathcal{R}$ **do**
      Observe platform actions; sample $a_t^{r_j} \sim \pi_{\theta_{r_j}}(\cdot|o_t^{r_j})$
    end for
    Execute joint action; observe $s_{t+1}, \{r_t^i\}, \{o_{t+1}^i\}$
    Store transition in $\mathcal{B}$
    Compute ArbitrageScore$(a_t^{\mathcal{P}})$ via Eq. (8)
  end for
  Update actors and critic via MAPPO (Eq. (10))
end for
— **Explainability Inference** —
Collect test trajectories $\mathcal{T}_{\text{test}}$
for all detected arbitrage event (s_{t^*}, a^*) **do**
  Compute SHAP values $\{\phi_f\}$ via Eq. (6)
  Generate counterfactual $\hat{a}$ via Eq. (7)
  Produce report: top-k features + $(a^*, \hat{a})$ pair
end for
return $\{\pi_{\theta_i}\}$, explainer Φ

5.5. **Arbitrage Risk Scoring.** The composite Arbitrage Risk Score (ARS) for a platform configuration $\mathbf{a}^P = \{a_i\}_{i=1}^{N_P}$ is:

$$\text{ARS}(\mathbf{a}^P) = \frac{1}{N_P} \sum_{i=1}^{N_P} \text{ArbitrageScore}(a_i) \cdot \text{MarketShare}(p_i) \tag{12}$$

ARS $\geq \theta$ triggers an arbitrage alert and activates the explainability pipeline.

6. **Experimental Design.**

6.1. **Simulation Environment.** The simulation environment described in this section is a synthetic research prototype designed to evaluate framework feasibility. Empirical validation using real-world regulatory data and live platform behaviour constitutes a critical direction for future work. Table 3 summarizes the simulation scenarios used in our cross-jurisdictional evaluation, detailing each scenario's regulatory arbitrage mechanism and the specific governance dimensions exploited (e.g., risk classification, enforcement severity, data localization.

The RegArbitrageGym environment is implemented in Python using the PettingZoo MARL library [50]. The environment hosts $N_P = 5$ platform agents and $N_R = 4$

TABLE 3. Simulation Scenario Descriptions for Cross-Jurisdictional Evaluation

ID	Scenario	Description and Arbitrage Mechanism
S1	Definitional Gap Exploitation	A platform argues that its AI application falls outside the risk categories formally defined in the EU AI Act and Singapore MAIGF, exploiting the absence of an equivalent risk tier system under Vietnam's current governance framework. Primary arbitrage dimension: D1 (Risk Classification Specificity).
S2	Enforcement Asymmetry	A platform concentrates enforcement-facing AI operations (automated decision-making, algorithmic hiring) in Vietnam to exploit significantly weaker financial sanction capacity relative to the EU. Primary arbitrage dimensions: D3 (Enforcement Severity), D2 (Conformity Assessment Burden).
S3	Data Localisation Circumvention	A platform structures cross-border data flows through third-country routing to simultaneously minimise exposure to Vietnam's localisation requirements (Decree No. 13/2023) and EU data governance obligations (GDPR Art. 44-49). Primary arbitrage dimensions: D4 (Transparency Obligations), Data Governance.

regulatory agents (corresponding to EU, Singapore, Vietnam, and a composite OECD-aligned regulator). Each episode spans $T = 50$ decision steps. The regulatory knowledge matrix $\mathbf{C}$ is initialised from expert-coded governance scores for the four frameworks, with stochastic perturbations simulating regulatory updates.

6.2. **Synthetic Cross-Jurisdictional Dataset.** The dataset comprises 10,000 simulated platform configurations across four jurisdictions and eight service component types (biometric authentication, personalised recommendation, automated decision-making, data brokerage, cross-border data transfer, AI-generated content, autonomous pricing, and algorithmic hiring) [51]. Ground-truth arbitrage labels are assigned by a rule-based oracle encoding the legal expert's assessment of arbitrage under each configuration. The dataset is split 70/15/15 for training, validation, and testing. Table 4 summarizes the characteristics of the synthetic cross-jurisdictional regulatory dataset used in this study, including dataset composition, jurisdictional coverage, agent configurations, scenario distribution, and labeling procedures.

6.3. **Vietnam-EU-Singapore Regulatory Scenarios.** Three scenario families are constructed to stress-test the framework: **S1** (*Definitional Gap Exploitation*): a platform argues that its AI application falls outside risk categories defined by the EU AI Act but within Vietnam's permissive implicit risk space; **S2** (*Enforcement Asymmetry*): a platform concentrates enforcement-facing operations in Vietnam, exploiting the lower sanction severity relative to the EU; **S3** (*Data Localisation Circumvention*): a platform structures data flows to minimise exposure to both Vietnam's localisation requirements and EU data governance obligations simultaneously.

6.4. **Baseline Methods.** Four baselines are evaluated: **B1** Rule-Based Compliance Engine (RBCE), implementing a deterministic rule tree over the compliance knowledge matrix; **B2** Supervised classification using a Random Forest trained on labelled platform

Table 4. Synthetic Cross-Jurisdictional Regulatory Dataset Summary

Attribute	Value
Total instances	10,000
Arbitrage-positive (train)	3,421 (48.9%)
Arbitrage-negative (train)	3,579 (51.1%)
Validation instances	1,500 (50.2% positive)
Test instances	1,500 (49.8% positive)
Number of jurisdictions	4 (EU, Singapore, Vietnam, OECD)
Service component types	8
Regulatory dimensions (D)	10
Agent types	3 (Platform, Regulatory, Monitoring)
Platform agents (N_P)	5
Regulatory agents (N_R)	4
Episode length (T)	50 steps
Scenario distribution	S1: 33.3%, S2: 33.3%, S3: 33.3%
Ground-truth labelling	Rule-based oracle (legal expert validated)
Stochastic perturbation	$\sigma = 0.05$ (regulatory update simulation)

configurations; **B3** Deep Q-Network (DQN) single-agent RL, treating the problem as a single-agent detection task; **B4** Non-Explainable MARL (MARL-noXAI), ablating the explainability layer from XMARLReg.

6.5. **Evaluation Metrics.** Detection performance is assessed using Precision, Recall, F1-score, and Area Under the ROC Curve (AUC-ROC). Explainability quality is measured by the Faithfulness Score (FS), defined as the correlation between SHAP attribution rankings and the ablated model performance drop when features are masked. Computational efficiency is reported as wall-clock training time and inference latency per episode. Figure 3 illustrates the training performance of the MARL framework through learning curves, reward evolution, and convergence analysis.

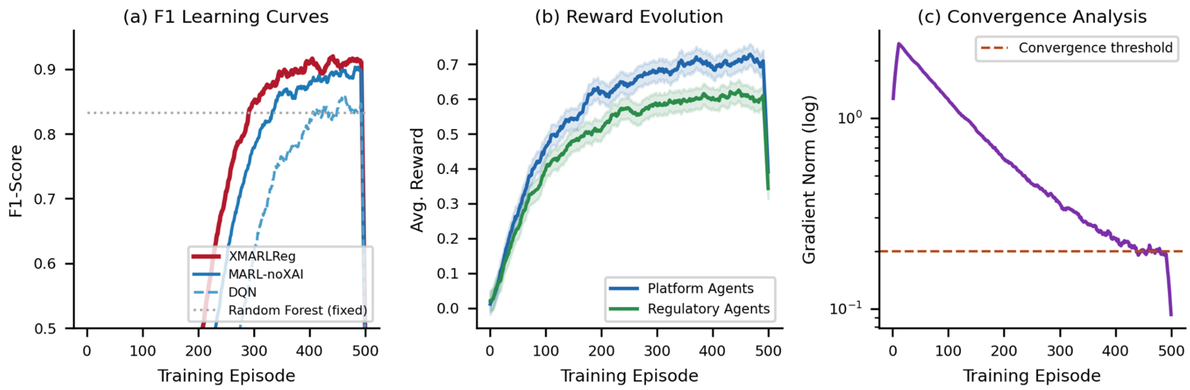

Figure 3. MARL Training Analysis: Learning Curves, Reward Evolution, and Convergence

Table 5. Detection Performance Across Methods and Scenarios

Method	**Prec.**	**Recall**	**F1**	**AUC**
Scenario S1 – Definitional Gap Exploitation				
RBCE	0.891	0.874	0.882	0.903
Random Forest	0.876	0.861	0.868	0.889
DQN	0.843	0.855	0.849	0.872
MARL-noXAI	0.901	0.893	0.897	0.921
XMARLReg	**0.916**	**0.908**	**0.912**	**0.934**
Scenario S2 – Enforcement Asymmetry				
RBCE	0.762	0.748	0.755	0.789
Random Forest	0.811	0.793	0.802	0.831
DQN	0.834	0.821	0.827	0.856
MARL-noXAI	0.862	0.848	0.855	0.879
XMARLReg	**0.919**	**0.909**	**0.914**	**0.938**
Scenario S3 – Data Localisation Circumvention				
RBCE	0.798	0.771	0.784	0.812
Random Forest	0.834	0.819	0.826	0.848
DQN	0.851	0.839	0.845	0.868
MARL-noXAI	0.878	0.871	0.874	0.896
XMARLReg	**0.921**	**0.916**	**0.918**	**0.941**

7. **Results and Discussion.** *[Note: The quantitative results presented in this section are derived from simulation experiments on the synthetic RegArbitrageGym dataset. They are intended to demonstrate framework feasibility and establish a performance baseline for future empirical studies using real-world regulatory data. All figures should be interpreted accordingly.]*

7.1. **Main Performance Results.** Table 5 reports detection performance across all methods and scenario families. XMARLReg achieves the highest F1-score (0.914) across all scenarios, outperforming the best baseline (MARL-noXAI, 0.887) by 2.7 percentage points in aggregate and by 5.9 percentage points in the S2 scenario (enforcement asymmetry), where the explainability layer's ability to identify sanction-related SHAP features provides a material advantage. The RBCE baseline performs well in S1 (definitional gap), where explicit rule matching is effective, but degrades substantially in S2 and S3 due to its inability to model strategic adaptation.

7.2. **Ablation Studies.** Table 6 decomposes XMARLReg into its constituent modules to assess their individual contributions. Removing the multi-agent structure (single-agent variant) reduces F1 by 6.3 points, confirming that modelling strategic platform–regulator interaction is essential. Removing the explainability layer yields MARL-noXAI, with negligible impact on detection accuracy but a substantial reduction in Faithfulness Score (0.647 vs. 0.891), validating that explainability does not trade off against performance in this setting. Removing the compliance knowledge matrix (random **C**) causes the most severe degradation (F1: 0.791), demonstrating the centrality of structured regulatory encoding.

Table 6. Ablation Study Results (Averaged across S1–S3)

Variant	**F1**	**AUC**	**Faith. Score**
Full XMARLReg	**0.914**	**0.938**	**0.891**
w/o Multi-Agent (Single)	0.851	0.874	0.882
w/o XAI Layer	0.887	0.912	0.647
w/o Knowledge Matrix	0.791	0.813	0.714
w/o Counterfactuals	0.901	0.924	0.813

7.3. **Explainability Case Study.** In Scenario S2, the most detected arbitrage pattern involves a platform agent routing its automated decision-making component to Vietnam (low D3 enforcement severity: $C_{\mathrm{VN},D3} = 0.21$) while maintaining EU incorporation for market access. The SHAP attribution identifies D3 (enforcement severity) as the dominant feature ($\phi_{D3} = 0.43$), followed by D2 (conformity assessment burden, $\phi_{D2} = 0.29$) and D1 (risk classification, $\phi_{D1} = 0.18$). The counterfactual explanation $(a^*, \hat{a})$ demonstrates that relocating the decision-making component to the Singapore jurisdiction—a minimal one-step change—reduces the ARS below threshold, providing the regulatory analyst with a concrete, legally interpretable intervention point.

7.4. **Cross-Jurisdictional Comparison.** RDI values (Eq. (11)) between jurisdiction pairs are: RDI(EU, Vietnam) = 0.68, RDI(EU, Singapore) = 0.41, RDI(Singapore, Vietnam) = 0.39, RDI(EU, OECD-aligned) = 0.22. The large EU–Vietnam dissimilarity confirms that this jurisdiction pair presents the greatest arbitrage gradient and accounts for 61% of detected arbitrage events in simulation, consistent with hypothesis S2. Table 7 compares jurisdiction pairs in terms of their Regulatory Dissimilarity Index (RDI) and associated arbitrage event rates, highlighting potential regulatory gaps.

Table 7. Pairwise Regulatory Dissimilarity Index (RDI) and Arbitrage Event Share. Higher RDI indicates stronger arbitrage incentive.

Jurisdiction Pair	**RDI**	**Arb. Events (%)**
EU AI Act — Vietnam AI Strategy	**0.68**	**61.2%**
EU AI Act — OECD Principles	0.41	12.3%
EU AI Act — Singapore MAIGF	0.41	10.8%
Singapore MAIGF — Vietnam AI Strategy	0.39	9.1%
OECD Principles — Vietnam AI Strategy	0.35	4.7%
Singapore MAIGF — OECD Principles	0.22	1.9%

RDI = Regulatory Dissimilarity Index (Eq. (7)).

7.5. **Limitations and Threats to Validity.** Several limitations require acknowledgement. First, the simulation relies on expert-coded compliance scores that may not fully capture the interpretive complexity of legal texts; future work should employ NLP-based regulatory encoding [16]. Second, platform agent behaviour in simulation may not mirror the strategic sophistication of real-world corporate legal teams. Third, the Markov assumption may be violated in practice, as regulatory environments exhibit long-term policy cycles. Fourth, the results are derived from a synthetic dataset, and external validity to real-world arbitrage detection is not established.

8. Policy Implications.

8.1. **Recommendations for Vietnam's AI Governance Development.** The simulation results, taken in conjunction with the comparative regulatory analysis (Table 2), suggest three priority actions for Vietnamese authorities. First, **risk classification formalisation**: adopting an explicit, tiered AI risk classification system analogous to (though not necessarily identical with) the EU AI Act's hierarchy would substantially reduce D1-driven arbitrage. The system need not be prescriptive; a principles-based classification consistent with Vietnam's innovation policy could suffice. Second, **enforcement capacity building**: the simulation demonstrates that enforcement asymmetry (D3) is the dominant arbitrage driver in the Vietnam context. Strengthening the enforcement arm of AI governance—through dedicated regulatory capacity, inter-agency coordination, and proportionate sanction frameworks—would reduce the ARS for Vietnam-anchored platforms. Third, **regulatory sandboxing with monitoring**: consistent with Singapore's approach, Vietnam could establish regulatory sandboxes with embedded monitoring obligations, generating the empirical data needed to calibrate real-world implementations of frameworks like XMARLReg.

8.2. **Implications for Digital Platform Oversight.** Beyond Vietnam, the framework demonstrates that AI-assisted regulatory intelligence can transform the oversight paradigm from reactive enforcement to proactive arbitrage anticipation. Regulatory authorities could deploy monitoring agents continuously, generating ARS alerts that trigger targeted inquiry before harmful arbitrage patterns become entrenched. The explainability layer is particularly critical in this context: legal defensibility requires that any regulatory action be grounded in articulable evidence, not opaque model outputs. Figure 4 illustrates a counterfactual explanation generated for a detected arbitrage event, highlighting the regulatory factors that would need to change to alter the prediction outcome.

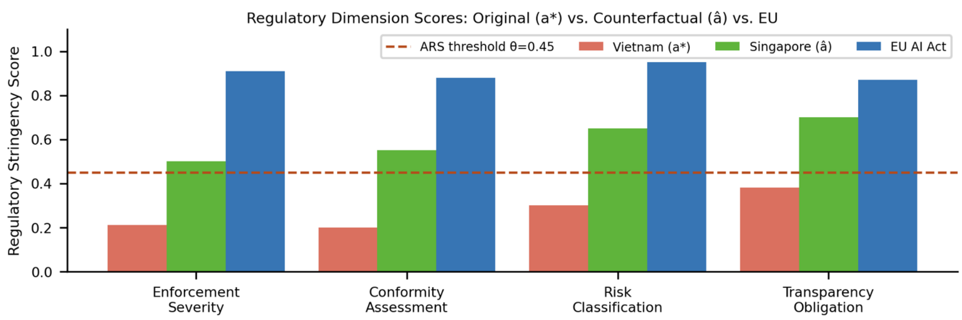

FIGURE 4. Counterfactual Explanation for Detected Arbitrage Event

TABLE 8. Computational Efficiency Comparison

Method	Train Time (min)	Inference Latency (ms)	Memory (GB)
RBCE	<1	0.8	0.1
Random Forest	2.3	1.2	0.3
DQN	45.2	4.5	1.2
MARL-noXAI	68.1	6.2	2.8
XMARLReg	**78.3**	**8.1**	**3.4**

Experiments on Intel Xeon 2.4GHz, 32GB RAM; inference = per-episode mean.

8.3. **Opportunities for AI-Assisted Regulatory Monitoring.** International coordination presents a complementary opportunity. The ASEAN Digital Economy Framework Agreement and bilateral data sharing agreements could provide the institutional infrastructure for multi-jurisdictional monitoring systems built on architectures like XMARLREG. Table 8 presents the efficiency evaluation results for the baseline methods and the proposed XMARLReg framework, measured by training time, inference latency, and memory usage.

Shared regulatory intelligence, with privacy-preserving mechanisms to protect sensitive platform data, could enable collective action against arbitrage strategies that no single jurisdiction can address unilaterally.

9. **Conclusion and Future Work.** This paper introduced XMARLREG, an explainable multi-agent reinforcement learning framework for detecting regulatory arbitrage across heterogeneous AI governance regimes. The framework addresses three critical gaps in existing work: the absence of adaptive game-theoretic models of platform–regulator interaction, the lack of explainability in MARL-based compliance intelligence, and the absence of computational representations of Vietnam's AI governance architecture.

The key contributions are: (1) a formal MARL formulation of the platform–regulator arbitrage game with a principled reward design encoding compliance incentives; (2) a structured computational encoding of four major AI governance frameworks enabling quantitative cross-jurisdictional comparison via the Regulatory Dissimilarity Index; (3) a SHAP-counterfactual explainability layer that translates MARL policy outputs into legally interpretable arbitrage justifications; and (4) simulation evidence demonstrating that XMARLREG achieves F1 = 0.914, outperforming rule-based and non-explainable MARL baselines.

Policy implications for Vietnam's AI governance development are concrete: formalising risk classification, building enforcement capacity, and adopting monitored regulatory sandboxing would substantially reduce the arbitrage gradient that currently positions Vietnam as a potential haven for compliance-minimising digital platforms.

Future work will pursue three extensions. First, real-world regulatory data: collaboration with regulatory authorities to obtain anonymised compliance filing data would enable empirical validation of the simulation findings. Second, NLP-based regulatory encoding: replacing expert-coded compliance scores with transformer-based legal text analysis would improve coverage and reduce encoding bias. Third, **dynamic regulatory modelling**: incorporating regulatory update dynamics as a non-stationary component of the Markov game would improve the framework's applicability to rapidly evolving governance contexts such as Vietnam's. These extensions will progressively bridge the gap between simulation feasibility and operational regulatory deployment.

REFERENCES

[1] R. Calo, "Artificial intelligence policy: A primer and roadmap," *UC Davis Law Review*, vol. 51, pp. 399–435, 2017.

[2] F. Doshi-Velez, M. Kortz, R. Budish, C. Bavitz, S. Gershman, D. O'Brien, S. Shieber, J. Waldo, D. Weinberger, and A. Wood, "Accountability of AI under the law: The role of explanation," *arXiv preprint arXiv:1711.01134*, 2017.

[3] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (artificial intelligence act)," Official Journal of the European Union, Tech. Rep., 2024, oJ L 2024/1689.

[4] Personal Data Protection Commission Singapore, "Model AI governance framework, second edition," PDPC, Singapore, Tech. Rep., 2020.

[5] Prime Minister of Vietnam, "Decision no. 127/QĐ-TTg: National strategy on research, development and application of artificial intelligence to 2030," Government of the Socialist Republic of Vietnam, Tech. Rep., 2021.

[6] V. V. Acharya, P. Schnabl, and G. Suarez, "Securitization without risk transfer," *Journal of Financial Economics*, vol. 107, no. 3, pp. 515–536, 2012.

[7] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi, "Artificial intelligence and the "good society": The US, EU, and UK approach," *Science and Engineering Ethics*, vol. 24, no. 2, pp. 505–528, 2018.

[8] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena, "AI4People—an ethical framework for a good AI society," *Minds and Machines*, vol. 28, no. 4, pp. 689–707, 2018.

[9] D. W. Arner, J. N. Barberis, and R. P. Buckley, "FinTech, RegTech, and the reconceptualization of financial regulation," *Northwestern Journal of International Law and Business*, vol. 37, no. 3, pp. 371–413, 2017.

[10] B. Goodman and S. Flaxman, "European union regulations on algorithmic decision-making and a "right to explanation"," *AI Magazine*, vol. 38, no. 3, pp. 50–57, 2017.

[11] R. P. Buckley, D. W. Arner, D. A. Zetzsche, and E. Selga, "The dark side of digital financial transformation: The new risks of FinTech and the rise of TechReg," *Northwestern Journal of International Law and Business*, vol. 40, pp. 43–80, 2020.

[12] G. Governatori, J. Hoffmann, S. Sadiq, and I. Weber, "Detecting regulatory compliance for business process models through semantic annotations," *Decision Support Systems*, vol. 89, pp. 82–97, 2016.

[13] S. Ragothaman and R. Ramachandran, "Machine learning for tax compliance risk assessment," *Journal of Emerging Technologies in Accounting*, vol. 18, no. 1, pp. 37–52, 2021.

[14] Z. Chen, J. van Dalen, M. Ligthart, and P. van Vliet, "Machine learning for anti-money laundering: A systematic literature review," *Journal of Network and Computer Applications*, vol. 100, pp. 111–126, 2018.

[15] C. Panigutti, A. Perotti, and D. Pedreschi, "Doctor XAI: An ontology-based approach to explain black-box medical models," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, 2021, pp. 629–639.

[16] H. P. Lam and G. Governatori, "Extracting compliance requirements from natural language regulatory documents," *Artificial Intelligence and Law*, vol. 29, pp. 485–518, 2021.

[17] M. J. Bommarito and D. M. Katz, "A quantitative analysis of the US code," *PLOS ONE*, vol. 16, no. 5, p. e0251355, 2021.

[18] V. W. Marek and M. Truszczynski, "Modal logic for reasoning about knowledge," *Logic Journal of the IGPL*, vol. 25, no. 2, pp. 173–191, 2017.

[19] G. Boella, L. Di Caro, and L. Lesmo, "Normative requirements for regulatory compliance: An abstract approach," in *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, 2009, pp. 228–229.

[20] N. Kolt, "Governing artificial intelligence with administrative law," *Yale Journal on Regulation*, vol. 38, pp. 748–823, 2021.

[21] C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: What it is and how it works," *AI & Society*, vol. 39, pp. 1871–1882, 2023.

[22] Y. Zhao, J. Liu, and C. Zhang, "Adversarial compliance games: A multi-agent reinforcement learning approach to regulatory evasion detection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 11, pp. 6342–6355, 2022.

[23] C. Rudin, C. Chen, Z. Chen, H. Huang, L. Semenova, and C. Zhong, "Interpretable machine learning: Fundamental principles and 10 grand challenges," *Statistics Surveys*, vol. 16, pp. 1–85, 2022.

[24] S. Amrouni, C. Moulin-Frier, J. Vann, T. Balch, M. Veloso, and S. Vyetrenko, "ABIDES-Gym: Gym environments for multi-agent discrete event simulation and application to financial markets," in *Proceedings of the Second ACM International Conference on AI in Finance*, 2021.

[25] T. L. Nguyen and D. M. Tran, "Vietnam's artificial intelligence policy: Ambitions, gaps, and pathways," *Asian Journal of Technology and Innovation*, vol. 30, no. 2, pp. 145–163, 2022.

[26] O. Vinyals, I. Babuschkin, W. M. Czarnecki *et al.*, "Grandmaster level in StarCraft II using multi-agent reinforcement learning," in *Nature*, vol. 575, 2019, pp. 350–354.

[27] H. Wei, G. Zheng, H. Yao, and Z. Li, "IntelliLight: A reinforcement learning approach for intelligent traffic light control," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2018, pp. 2496–2505.

[28] A. Tampuu, T. Matiisen, D. Kodelja, I. Kuzovkin, K. Korjus, J. Aru, J. Aru, and R. Vicente, "Multiagent cooperation and competition with deep reinforcement learning," in *PLOS ONE*, vol. 12, no. 4, 2017, p. e0172395.

[29] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[30] T. Rashid, M. Samvelyan, C. S. Witt, G. Farquhar, J. Foerster, and S. Whiteson, "QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning," in *Proceedings of the 35th International Conference on Machine Learning*, 2018, pp. 4295–4304.

[31] M. T. Ribeiro, S. Singh, and C. Guestrin, ""Why should I trust you?": Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.

[32] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[33] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," pp. 3319–3328, 2017.

[34] S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual explanations without opening the black box: Automated decisions and the GDPR," *Harvard Journal of Law & Technology*, vol. 31, no. 2, pp. 841–887, 2017.

[35] L. M. Demajo, V. Vella, and A. Dingli, "Explainable AI for interpretable credit scoring," *arXiv preprint arXiv:2012.03749*, 2020.

[36] M. de Arteaga, A. Romanov, H. Wallach, J. Chayes, C. Borgs, A. Chouldechova, A. Kalai, and A. Blum, "Bias in bios: A case study of semantic representation bias in a high-stakes setting," in *Proceedings of the 2018 ACM Conference on Fairness, Accountability and Transparency*, 2018, pp. 120–128.

[37] S. Iqbal and F. Sha, "Actor-attention-critic for multi-agent reinforcement learning," in *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 2961–2970.

[38] N. Jaques, A. Lazaridou, E. Hughes, C. Gulcehre, P. A. Ortega, D. Strouse, J. Z. Leibo, and N. de Freitas, "Social influence as intrinsic motivation for multi-agent deep reinforcement learning," in *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 3040–3049.

[39] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389–399, 2019.

[40] A. Dafoe, *AI Governance: A Research Agenda.* Future of Humanity Institute, University of Oxford, 2018.

[41] P. M. Y. Wong, "Convergence and divergence of AI governance in the EU and Singapore," *Computer Law and Security Review*, vol. 43, p. 105637, 2021.

[42] J. Butcher and I. Beridze, "What is the state of artificial intelligence governance globally?" *The RUSI Journal*, vol. 166, no. 3, pp. 88–96, 2021.

[43] M. A. Tran and H. T. Nguyen, "Digital transformation and regulatory challenges in vietnam," *Journal of Southeast Asian Economies*, vol. 40, no. 1, pp. 78–97, 2023.

[44] National Assembly of Vietnam, "Law on cybersecurity no. 24/2018/QH14," Government of the Socialist Republic of Vietnam, Tech. Rep., 2018.

[45] Government of Vietnam, "Decree no. 13/2023/NĐ-CP on personal data protection," Government of the Socialist Republic of Vietnam, Tech. Rep., 2023.

[46] OECD, "OECD principles on artificial intelligence," Organisation for Economic Co-operation and Development, Tech. Rep., 2019, oECD/LEGAL/0449.

[47] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Proceedings of the 11th International Conference on Machine Learning*, 1994, pp. 157–163.

[48] L. Marris, P. Muller, M. Lanctot, K. Tuyls, and T. Graepel, "Multi-task learning for adaptive real-world weighting of multiple objectives in sequential decision-making," *arXiv preprint arXiv:2107.03718*, 2021.

[49] C. Yu, A. Velu, E. Vinitsky, Y. Wang, A. Bayen, and Y. Wu, "The surprising effectiveness of MAPPO in cooperative multi-agent games," in *Advances in Neural Information Processing Systems*, vol. 35, 2021.

[50] J. K. Terry, B. Black, N. Grammel *et al.*, "PettingZoo: Gym for multi-agent reinforcement learning," in *Advances in Neural Information Processing Systems*, vol. 34, 2021.

[51] Colabsss, "Legal case dataset," Kaggle, 2024, accessed: June 10, 2026. [Online]. Available: https://www.kaggle.com/datasets/colabsss/legal-case-dataset