



Constrained Adaptive Negotiation Agent of ODRL Offers Under Incomplete and Asymmetric Information

Iason Sotiropoulos^{1,2,*}, Ioannis Karvelas^{1,†}, Elissaios Sarmas², Stamatia Rizou¹, and Vangelis Marinakis²

¹ SingularLogic Single Member S.A., Cholargos, Athens, Greece

² School of Electr. & Comp. Engineering, National Technical University of Athens, Zografou, Athens, Greece

Abstract

The Open Digital Rights Language (ODRL) has emerged as a key policy language for governing data usage in data spaces and decentralized digital ecosystems. While ODRL enables machine-readable representation and enforcement of permissions, prohibitions, and obligations, automated negotiation under ODRL constraints remains insufficiently addressed. Autonomous agents should uphold these constraints on top of negotiating heterogeneous policy terms under incomplete information. This paper presents a policy-aware negotiation agent that operates directly on ODRL policy constraints while ensuring the semantic interpretability of negotiated offers. This work focuses on constraints associated with permissions, allowing multiple actions to be negotiated in parallel, while prohibitions and duties are considered fixed. The proposed framework combines per-issue Bayesian belief models capable of handling mixed-type policy constraints, strict enforcement of non-negotiable constraints, and adaptation mechanisms that guide acceptance decisions and counter-offer generation. Unlike utility-driven or reinforcement learning approaches, the proposed agent enforces non-negotiable constraints by construction and restricts exploration to the feasible policy space. We position this framework as a practical step toward regulating autonomous agents with ODRL in open systems, such as data spaces, and report implementation insights from translating ODRL constraint semantics into executable negotiation logic.

1 Introduction

Data sharing has become a central challenge in the rapidly evolving digital landscape, driven by the increasing availability of large and heterogeneous data sources and the demand of modern predictive and analytics systems. This growth has accelerated the development of advanced infrastructures and architectures for storing, managing, and disseminating data across organizational boundaries. Beyond purely technological advancements, significant attention has been directed toward issues of data sovereignty, trust, and control. In particular, concerns related to

*Corresponding author

†Corresponding author

interoperability, secure data exchange, and the emergence of data marketplaces highlight the need for solutions that balance accessibility with the protection of data ownership and usage rights. In parallel, regulatory frameworks, such as the General Data Protection Regulation (GDPR) and emerging data governance acts, have further emphasized the importance of accountability, transparency, and compliance in data sharing processes, reinforcing the need for technically enforceable governance mechanisms. The convergence of these multidimensional frameworks and policies, together with the sophisticated technological elements of data sharing and hosting, creates a complex landscape — one shaped by a wide range of constraints as well as opportunities.

In the evolving landscape of data exchange, data spaces have emerged as a key paradigm for enabling trusted and secure data sharing while preserving data sovereignty. Ecosystems, such as the International Data Spaces Association (IDSA),¹ GAIA-X,² and Eclipse Data-space Components (EDC),³ support data governance, sovereignty, and interoperability while enabling controlled data sharing under predefined policies and legal frameworks. These initiatives promote decentralized and federated architectures in which data providers retain control over their assets, avoiding the limitations of centralized data platforms. Since these ecosystems involve diverse participants, uniform semantic interpretation and processing become essential for data exchange in self-regulated environments [5]. Without shared semantics, differences in data models, formats, and contextual assumptions can lead to misinterpretation, reduced data quality, and limited interoperability across participants.

Semantically annotated resources enable consistent interpretation of data regardless of origin, making them discoverable, accessible, interoperable, and reusable [4]. Such annotations typically rely on shared vocabularies and ontologies, which provide formal and machine-interpretable descriptions of data and its context. This formalization is particularly important in domains such as urban ecosystems, where heterogeneous stakeholders depend on interoperable data governance frameworks [21]. In these contexts, data originate from multiple sources, including sensors, public administrations, and private organizations, further increasing the need for harmonized semantic representations. Moreover, the complexity of managing data usage rights and ownership increases the need for mechanisms that support contract negotiation between data providers and consumers in multi-party settings [8]. Such negotiations may involve dynamic constraints, including temporal restrictions, purpose limitations, and obligations that must be fulfilled during or after data usage. The shift toward seamless data exchange and automated negotiation in policy-constrained environments creates a strong need for machine-readable and enforceable usage policies.

Traditional data processing agreements, often expressed as “Terms and Conditions”, are coarse-grained and difficult for software systems to interpret and enforce automatically. These agreements are typically written in natural language, introducing ambiguity and requiring manual interpretation, limiting their potential for application in automated systems. Consequently, data owners must describe usage policies in machine-processable languages [13]. This enables not only automated enforcement but also validation, monitoring, and auditing of data usage in distributed environments. The Open Digital Rights Language (ODRL),⁴ a W3C-standard ontology for expressing usage policies, provides a formal vocabulary for specifying permissions, prohibitions, and obligations associated with data usage [7]. Its extensible information model allows domain-specific adaptations, making it suitable for a wide range of application scenarios. ODRL is well suited for formalizing usage control patterns in data spaces, with extensions

¹<https://internationaldataspaces.org/>

²<https://gaia-x.eu/>

³<https://projects.eclipse.org/projects/technology.edc>

⁴<https://www.w3.org/TR/odrl-model/>

already proposed in various domains [8]. In addition, it is increasingly used to define policies within authentication and authorization frameworks, enabling fine-grained access control in data spaces [19]. Despite these formalizations, challenges remain in aligning semantic data descriptions with policy representations and integrating them into automated negotiation and enforcement workflows, particularly in dynamic and multi-stakeholder environments.

1.1 Related Work

Despite the advantages that formalizations like the ODRL provide, negotiating over its policies remains challenging. Policy offers often involve mixed-type attributes and constraints that complicate automated reasoning and negotiation. These challenges become more pronounced in dynamic contexts such as smart cities, where numerous stakeholders (e.g., public authorities, private companies, infrastructure providers, citizens, etc.) produce and consume large volumes of heterogeneous data [12]. Existing automated negotiation systems are rarely designed to simultaneously handle heterogeneous issue types, adaptive learning requirements, and strict policy compliance [11]. The problem becomes even more complex when agents must both optimize negotiation outcomes and ensure that agreements remain consistent with predefined usage constraints [18]. Although automated negotiation has been widely studied, little work focuses on negotiation directly over ODRL policies, particularly under incomplete and asymmetric information.

Automated negotiation research has evolved from classical utility-based strategies to more advanced AI-driven approaches. Traditional methods rely on predefined utility functions and concession strategies within structured negotiation spaces [1]. Opponent modeling techniques, such as Bayesian learning, attempt to infer hidden preferences from limited observations, while argumentation-based negotiation introduces structured exchanges of justifications [6, 20]. Game-theoretic models provide equilibrium solutions but often assume complete information and perfect rationality, which rarely hold in real-world data-sharing scenarios [9, 15]. More recently, Reinforcement Learning (RL) and Large Language Models (LLMs) have enabled adaptive and language-capable negotiation agents [23, 14]. However, RL-based approaches typically aim to maximize rewards in dynamic environments, which may lead to agreements that inadvertently violate policy constraints [2, 17, 16], while LLM-based systems struggle to balance flexibility with strict adherence to machine-readable constraints [3, 22]. To date, none explicitly supports negotiation over formal ODRL policy semantics. The closest effort is MINiDM [10], which models negotiation in data marketplaces as a multi-agent, multi-issue bargaining problem and guarantees convergence to Nash equilibria and Pareto-optimal outcomes in convex strategy spaces. However, it assumes fully negotiable issues and does not consider formal policy semantics, non-compensable constraints, or probabilistic uncertainty about opponent preferences.

1.2 Motivation

The limitations of automated negotiation research highlight the need for negotiation strategies capable of handling mixed-type issues under incomplete and asymmetric information while ensuring strict policy compliance. To address this gap, this paper introduces a novel negotiation framework tailored to policy-governed data spaces. The proposed approach integrates per-issue probabilistic beliefs, a policy-aware offer space, and adaptive control mechanisms that promote stable convergence. This enables the negotiation agent to reason under uncertainty while ensuring that every generated offer remains compliant with formal ODRL policies. By balancing learned beliefs about an opponent’s acceptable constraints with strict policy enforcement, the

framework bridges the gap between adaptive optimization and formal policy adherence, enabling reliable automated negotiation in complex data-sharing environments. In the current formulation, negotiation is performed over the constraints associated with ODRL permissions, while multiple actions can be negotiated in parallel through independent issue sets. Prohibitions and duties are currently treated as fixed, non-negotiable policy elements and are left for future work.

2 Algorithmic Overview

At each round, the agent (i) parses an incoming ODRL offer into internal issues, (ii) updates adaptive control signals and learns per-issue Bayesian beliefs from feasible observations, (iii) decides acceptance using belief-based confidence and semantic compatibility, and (iv) if not accepted, produces a counter-offer by sampling from beliefs within the feasible region. The core idea is that learning and proposal generation never leave the policy-feasible space. The remainder of this section introduces the formal framing of the negotiation problem and the design of the proposed autonomous agent.

We consider a bilateral negotiation between two agents over a finite set of ODRL actions $\mathcal{A}$. Each action $a \in \mathcal{A}$ contains a set of permissions that are negotiated independently through a set of issues. For a fixed action a , let $\mathcal{I}_a = \{1, \dots, n_a\}$ denote its associated issues. In the current formulation, negotiation is restricted to constraints associated with ODRL permissions. This enables multiple actions to be negotiated in parallel through independent issue sets. Prohibitions and duties, however, are treated as fixed and non-negotiable policy elements and are therefore outside the present negotiation space. Each issue $i \in \mathcal{I}_a$ corresponds to a single ODRL constraint triple $(\ell_i, \text{op}_i, r_i)$, where ℓ_i is the leftOperand, op_i is the operator, and r_i is the rightOperand. Additionally, each issue is associated with a type $T_i \in \{\text{numeric}, \text{enumeration}, \text{set}\}$, determined by the type of the initial offer's r_i . Each type is equipped with an internal representation map $\varphi_i : \mathcal{D}_i \rightarrow \mathcal{X}_i$, with inverse φ_i^{-1} , where $\mathcal{D}_i$ denotes the original value domain and $\mathcal{X}_i$ the corresponding internal representation space. For convenience, we denote by $x_{i,t} = \varphi_i(r_{i,t})$ the internal representation of the agent's most recent values. For each action a , the agent maintains a set of non-negotiable ODRL-style constraints $\mathcal{H}_a$. Each issue $i \in \mathcal{I}_a$ may be associated with multiple hard constraints indexed by j . Formally,

$$\mathcal{H}_a = \{H_{i,j} \mid i \in \mathcal{I}_a, j = 1, \dots, m_i\}_a, \quad (1)$$

where m_i denotes the number of hard constraints defined for issue i . Each constraint is compiled into an executable predicate

$$H_{i,j}(\text{op}, x) = \mathbb{I}[\text{Sat}(\text{op}_{i,j}^H, x_{i,j}^H; \text{op}, x) = 1], \quad x_{i,j}^H = \varphi_i(r_{i,j}^H). \quad (2)$$

The agent operates under incomplete information, in the notion that the opponent's true acceptance region for each issue is unknown and inferred online from observed offers. To this end, the agent maintains, for every issue $i \in \mathcal{I}_a$, a belief model $\mathcal{B}_i$ that approximates the opponent's preference over the transformed issue domain $\mathcal{X}_i$. Each belief is updated after receiving an offer (learning phase), provided that the incoming offer satisfies the respected issue's hard rules $H_{i,j}(\text{op}, x)$, and queried during acceptance checking (acceptance phase) and counter-offer construction (sampling phase). Since issue domains are heterogeneous, the agent uses a type-specific family of distributions that matches the geometry of each issue, namely a Beta distribution for scalar numeric/temporal issues, a Dirichlet model for categorical (enumeration)

issues, and independent Beta–Bernoulli factors for set-valued issues:

$$\begin{aligned}
\text{Numeric Issues: } z &= \psi_i(x) \in [0, 1], \quad z \sim \text{Beta}(\alpha_i, \beta_i), \\
\text{Enumeration Issues: } \theta_i &\sim \text{Dirichlet}(\boldsymbol{\alpha}_i) \text{ over labels } \mathcal{L}_i, \\
\text{Set Issues: } \pi_{i,e} &\sim \text{Beta}(\alpha_{i,e}, \beta_{i,e}) \text{ independently for } e \in \mathcal{U}_i.
\end{aligned} \tag{3}$$

This choice of distributions reflects the structural properties of each issue type, allowing lightweight online updates while preserving interpretability and efficient sampling.

Rather than relying on a fixed concession curve or pure utility maximization, the agent adapts its behavior at each round using two alignment signals derived from operator-aware issue similarities $\sigma_i(t) \in [0, 1]$:

$$c_a(t) = \frac{1}{|\mathcal{I}_a|} \sum_i i \in \mathcal{I}_a \sigma_i(t), \quad s_a(t) = 0.7 \max_i i \in \mathcal{I}_a (1 - \sigma_i(t)) + 0.3, Q0.95(1 - \sigma_i(t)). \tag{4}$$

The weighting coefficients in Eq. 4 are selected to prioritize worst-case conflicts while maintaining robustness to noise. In particular, the maximum term ensures that critical disagreements dominate the severity signal, while the high-quantile component captures near-worst-case behavior without being overly sensitive to outliers. The choice of the 95th percentile follows standard practice for high-quantile estimation, and the relative weighting (0.7/0.3) was determined through preliminary calibration runs and general design principles emphasizing bounded adaptation and sensitivity to critical conflicts.

Intuitively, $c_a(t)$ captures the overall level of agreement across issues, while $s_a(t)$ emphasizes persistent disagreement by focusing on the most conflicting aspects of the negotiation. The use of the maximum term ensures that a single critical conflict is not ignored, whereas the high-quantile term provides robustness against noise and prevents overreaction to isolated deviations.

For each issue i , the similarity $\sigma_i(t)$ provides a normalized measure of agreement between the agent’s and the opponent’s positions at round t , taking into account both value differences and operator semantics. For numeric issues, $\sigma_i(t)$ is computed based on the normalized distance between values, adjusted by operator compatibility (e.g., directional thresholds), such that closer and semantically consistent values yield higher similarity. For set and enumeration issues, $\sigma_i(t)$ reflects the degree of overlap and mutual satisfaction, combining structural similarity (e.g., Jaccard overlap) with operator-specific fulfillment conditions (e.g., inclusion, exclusion, or equality constraints). In all cases, values close to 1 indicate strong agreement, while values close to 0 indicate conflict.

These per-issue similarities enable the agent to distinguish aligned from conflicting issues and to concentrate concessions where disagreement is most pronounced, resulting in targeted and semantically consistent negotiation behavior. At the aggregate level, $c_a(t)$ captures global alignment across issues, while $s_a(t)$ ensures that a small number of highly conflicting issues cannot be masked by a high average similarity. Together, these signals drive the adaptive modulation of the agent’s control parameters at each round, including the learning rate, inertia for discrete and combinatorial issues, and parameters governing set sampling and plausibility. Adaptation also tunes acceptance aggressiveness through an effective risk tolerance:

$$g_a(t) = c_a(t)(1 - s_a(t)), \quad \rho_{\text{eff}}(t) = \rho + g_a(t)(\rho_{\max} - \rho), \tag{5}$$

where $\rho \in (0, 1)$ represents a nominal risk and $\rho_{\max} < 1$ an upper bound, so that when negotiation is globally aligned (high c_a , low s_a) the agent becomes more willing to accept borderline values to close agreement, while remaining conservative when incompatibility persists.

Negotiation Loop. At each round t , for each action $a \in \mathcal{A}$, the agent receives the opponent’s ODRL offer $o_{a,t}$ (the offer corresponding to action a at round t) and internalizes it into typed issue representations $(op_{i,t}, x_{i,t})$, where values from the ODRL domain D_i are mapped to a computational domain X_i . It first evaluates operator-aware similarities $\sigma_i(t)$ across issues and aggregates them into the action-level context and severity scores, $c_a(t)$ and $s_a(t)$, which quantify global alignment and persistent disagreement. These signals are then used to adapt the action-specific control parameters, through bounded updates, ensuring stable and responsive behavior. Belief learning is subsequently performed on each issue under strict feasibility control. For every issue i , the agent verifies whether the incoming constraint satisfies all hard predicates $H_{i,j}$. If a violation occurs, the belief $\mathcal{B}_i$ is not updated and the position is deterministically projected back to the feasible region defined by the hard constraints. Otherwise, the belief is updated using operator-aware Bayesian rules appropriate to the issue type, controlled by the adapted parameters. In particular, Beta posterior updates are applied for numeric issues, Dirichlet concentration updates for enumeration issues, and Beta–Bernoulli evidence increments for set-valued issues. The magnitude of these updates is governed by the learning rate, which is dynamically adjusted based on contextual alignment and volatility signals, ensuring controlled responsiveness to opponent behavior. Next, acceptance is evaluated on each issue by combining deterministic semantic compatibility with probabilistic confidence derived directly from the updated belief model. For each issue $i \in \mathcal{I}_+$, the confidence score $\text{conf}_i(t)$ is computed from the corresponding posterior distribution (e.g., Beta tail probabilities, Dirichlet mass over proposed labels, or inclusion probabilities under Beta–Bernoulli factors). An issue is accepted whenever Eq. 6 holds true and all hard constraints are satisfied.

$$\text{Accept}_i(t) = \mathbb{I}[\text{conf}_i(t) \geq 1 - \rho_{\text{eff}}(t)] \quad (6)$$

In simple terms, the agent accepts an issue when its observed offer is sufficiently likely under its learned belief about the opponent, adjusted by a dynamically adapted risk tolerance. If all issues of an action are accepted, an ODRL agreement is returned. Otherwise, a counter-offer is generated by sampling from the updated beliefs—using truncated Beta sampling for numeric issues, Dirichlet sampling with inertia for enumeration, and threshold-based selection with cardinality control for sets—followed by feasibility projection and mapping back to ODRL form. This ensures policy-compliant, belief-guided convergence.

3 Results

This section presents an empirical evaluation of the proposed negotiation framework, focusing on convergence behavior and outcome stability. The approach was evaluated through 10,000 independent runs of the same bilateral negotiation scenario between a provider (initial offer) and a consumer (counter-offers) under identical initial conditions. Table 1a shows the initial offers, while Table 1b summarizes the non-negotiable constraints that define the feasible negotiation domain. This configuration ensures guaranteed feasibility while maintaining strategically challenging initial conditions, allowing convergence behavior to reflect the negotiation mechanism’s ability to resolve conflicts and exploit the feasible region.

All experiments were conducted on a standard high-performance workstation (Intel i9-class CPU, 64GB RAM), and the implementation was executed in Python⁵. In each run, both agents start from the same initial offers and are subject to the same hard constraints. The

⁵The full code of the implementation, alongside with the experiments are available at: <https://github.com/isotiropoulos/BayesianNegotiator>

only variability arises from stochastic belief sampling and adaptive exploration. Negotiations proceed in alternating rounds until agreement or infeasibility is detected. A negotiation is considered successful when an offer is accepted consecutively by both agents, indicating mutual agreement on the proposed terms.

Table 1: Initial offers and hard constraints for provider and consumer agents used in all 10,000 negotiation runs. The `delayPeriod` constraint follows the ISO 8601 format.

(a) Initial offers

Constraint	Type	Provider	Consumer
<code>percentage</code>	Numeric	≤ 80	≥ 70
<code>count</code>	Numeric	≤ 1500	≥ 1700
<code>delayPeriod</code>	Duration	<code>= PT15M</code>	$\geq$ <code>PT20M</code>
<code>absoluteTemporalPosition</code>	Date	$\leq$ <code>01-07-2025</code>	$\geq$ <code>01-06-2025</code>
<code>fileFormat</code>	Enumeration	<code>= JSON</code>	<code>= XML</code>
<code>language</code>	Set	<code>isAnyOf{en, fr, gr}</code>	<code>= en</code>
<code>industry</code>	Enumeration	<code>= research</code>	<code>isAllOf{research, production}</code>
<code>purpose</code>	Set	<code>isAllOf{research, analytics, open-source}</code>	<code>isAnyOf{analytics, publication}</code>

(b) Hard constraints

Constraint	Provider	Consumer
<code>percentage</code>	≤ 90	≥ 65
<code>count</code>	≤ 1800	≥ 1300
<code>delayPeriod</code>	$\leq$ <code>PT30M</code>	$\geq$ <code>PT15M</code>
<code>absoluteTemporalPosition</code>	$\leq$ <code>31-08-2025</code>	$\geq$ <code>30-05-2025</code>
<code>fileFormat</code>	<code>= JSON</code>	<code>isAnyOf{JSON, XML}</code>
<code>language</code>	<code>isAnyOf{en, fr, gr}</code>	<code>= en</code>
<code>industry</code>	<code>isAnyOf{research, experimentation}</code>	<code>isAnyOf{research, production}</code>
<code>purpose</code>	<code>isAnyOf{research, analytics}</code>	<code>isAnyOf{analytics, publication}</code>

3.1 Convergence and Computational Cost of Counter-Offer Generation

Across all runs, 99.02% of negotiations terminate successfully with a mutually accepted agreement, while the remaining cases end due to infeasibility. Convergence is rapid: over 90% of successful negotiations conclude by round 7 and more than 99% by round 11. This is especially remarkable considering that the provider permits a maximum of 40 rounds and does not allow agreement before round 3, whereas the consumer permits up to 80 rounds and does not allow agreement before round 5. Agreements beyond this point are rare, indicating that unresolved

negotiations are typically infeasible rather than slow to converge. This demonstrates that the adaptive mechanisms of the framework effectively promote early convergence.

Counter-offer generation incurs higher computational cost in the initial rounds due to belief initialization and exploratory sampling. In early rounds, the provider requires on average about 17 ms per counter-offer, while the consumer peaks at roughly 20 ms. As posterior beliefs stabilize, computation time decreases rapidly. By approximately round 6, both agents stabilize at around 4–6 ms per round. These results confirm that the framework remains computationally lightweight even with belief sampling and Monte Carlo-based evaluation, supporting real-time and large-scale deployment.

3.2 Distributions of Agreed Values

Out of the 10,000 negotiation runs, the mechanism reaches agreement in 9,902 cases. The following analysis therefore considers only successful negotiations. Figure 1 summarizes the distributions of final agreements for numeric and temporal issues using box plots, highlighting dispersion, central tendency, and potential outliers across runs. The close alignment between

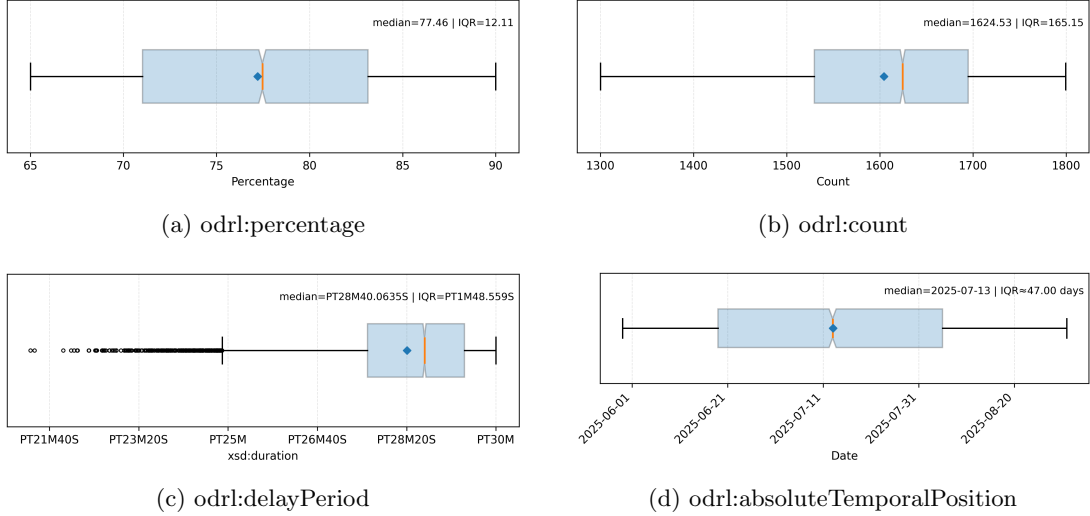


Figure 1: Final distributions of agreed values obtained across 10,000 negotiation runs. Boxes indicate the interquartile range (IQR), the central notch marks the median, whiskers denote the full observed range, and diamond markers show the mean.

mean and median values indicates that agreements consistently concentrate around a common region of the outcome space rather than being driven by extreme offers. At the same time, the non-zero interquartile ranges show that the process maintains limited variability, reflecting exploration during negotiation while still converging reliably to stable outcomes.

For **percentage** (Figure 1a), the median agreement is 77.46 with an interquartile range (IQR) of 12.11. The distribution is compact and approximately symmetric, indicating stable convergence to a narrow region of mutually acceptable values. A similar pattern appears for **count** (Figure 1b), where the median is 1624.53 with an IQR of 165.15, showing limited dispersion despite the larger numeric scale.

For **delayPeriod** (Figure 1c), the median agreement is approximately PT28M40S with an IQR of PT1M49S. This tight clustering is notable given the initially conflicting offers (= PT15M vs.

$\geq \text{PT20M}$). Similarly, **absoluteTemporalPosition** (Figure 1d) converges around mid-July 2025 with an IQR of roughly 47 days, indicating consistent compromise within the feasible temporal range.

Enumeration and set-valued issues exhibit deterministic outcomes. For **fileFormat**, **language**, **industry**, and **purpose**, all successful negotiations converge to a single agreement (dominant share= 100%, entropy=0). Specifically, **fileFormat** resolves to **JSON** due to the provider’s constraint, while **language** converges to **en** following the consumer’s requirement. The remaining issues consistently converge to **research** (**industry**) and **analytics** (**purpose**), reflecting the interaction between constraint overlap and negotiation dynamics.

Overall, the results demonstrate that the proposed mechanism produces robust agreements across heterogeneous issue types. Numeric and temporal issues converge to tightly concentrated values without multimodal behavior, while categorical issues produce deterministic and policy-compliant outcomes. Across 9,902 successful negotiations, the framework consistently balances exploration and stability, yielding reliable agreements despite stochastic sampling and diverse issue representations.

4 Limitations

While the proposed framework demonstrates strong empirical performance, several limitations should be acknowledged. First, the current evaluation is based on a single negotiation scenario with fixed initial conditions and constraint structures, and follows a self-play design where both agents implement the same decision mechanism. Although this enables controlled analysis of convergence behavior, it does not capture the diversity of real-world negotiation settings, nor does it account for heterogeneous or adversarial strategies. In addition, the absence of established open implementations of negotiation agents operating directly on ODRL policy constraints limits the ability to perform direct experimental comparisons. As a result, the evaluation focuses on internal performance characteristics rather than benchmarking against standardized baselines. Future work should include a broader range of scenarios and the development of shared benchmarks for ODRL-based negotiation. Second, the proposed approach focuses on bilateral negotiation and does not currently support multi-party settings. Extending the framework to handle more than two agents, including coalition formation and multi-lateral constraint interactions, remains an open challenge. Third, convergence is not guaranteed in the general case. As the framework relies on stochastic Bayesian belief updates and adaptive sampling, agreement depends on the existence of a non-empty feasible intersection and on the stabilization of posterior beliefs and control parameters. While empirical results show rapid convergence in the evaluated scenario, adversarial dynamics or persistent uncertainty may lead to non-convergent behavior. Finally, the current formulation is restricted to negotiation over constraints associated with ODRL permissions. Prohibitions and duties are treated as fixed, non-negotiable policy elements and are therefore outside the present negotiation space. Extending negotiation to include these components introduces additional semantic and logical complexity, which is left for future work.

5 Discussion and Future Work

In this paper, we proposed a novel framework for automated negotiations over ODRL policies by providing the algorithmic approach that enables optimal negotiation outcomes based on the constraints imposed by the involved parties (provider and consumer). The framework

was evaluated by conducting an experiment of multiple computational iterations over a fixed bilateral negotiation between two parties. The empirical results of the framework testing demonstrate high convergence reliability, rapid stabilization, and robust outcome concentration under stochastic exploration. The 99.02% success rate confirms that the negotiation dynamics effectively exploit the feasible region defined by the hard constraints. The small fraction of non-successful runs (0.98%) reflects rare stochastic trajectories reaching the round limit rather than structural incompatibility. Preliminary inspection suggests that these cases do not appear to stem from structural infeasibility or oscillatory behavior but rather from stochastic stagnation under finite round limits, where negotiation remains in a narrow region of the feasible space without reaching mutual acceptance. However, a detailed cycle analysis remains future work. Nevertheless, no persistent or systematic oscillatory patterns were observed across the evaluated runs, indicating that adaptive belief updates combined with feasibility enforcement tend to guide the negotiation toward stable regions of the solution space.

The analysis of the distribution of the agreed values, further confirms robustness across heterogeneous issue types. Numerical issue agreements are compact and unimodal, with close mean–median alignment and limited dispersion, indicating stable convergence without extreme outliers. At the same time, non-zero interquartile ranges confirm that the mechanism does not collapse deterministically, but preserves bounded variability driven by exploration. Categorical and set-valued issues converge deterministically in all successful runs. This behavior is structurally explained by the interaction between overlapping hard constraints and posterior concentration in the Dirichlet and Beta–Bernoulli models, which rapidly amplify the single semantically admissible solution.

Furthermore, we note that the convergence is rapid, with more than 90% of agreements are reached by round 7 and almost all by round 11. This behavior highlights the role of the context–severity mechanism, which reduces exploratory variance once semantic alignment is detected. Early stochastic sampling allows exploration of the feasible region, while adaptive cooling progressively narrows proposals toward mutually acceptable subspaces. Furthermore, from a computational perspective, the stabilized per-round cost of 4–6 ms confirms that the framework remains computationally lightweight, making it suitable for real-life applications. Notably, the decrease in counter-offer generation time additionally reflects a posterior stabilization, in the sense that as beliefs concentrate, fewer exploratory samples are required for confident acceptance decisions.

Overall, the results demonstrate that the proposed mechanism achieves a balanced trade-off between exploration and stability, consistently converging to policy-compliant and robust agreements across runs. This suggests that the solution can be scalable and suitable for integration in real world scenarios. Building on the results presented, future work will focus on extending both the evaluation and the applicability of the proposed framework. In particular, the current self-play setting will be expanded to include heterogeneous negotiation agents with diverse strategies, enabling a more realistic assessment under asymmetric and potentially adversarial conditions. At the same time, there is a clear need for the development of standardized benchmark scenarios for ODRL-based negotiation, allowing for reproducibility and meaningful comparison across different approaches. From a modeling perspective, the scope of negotiation will be broadened to include additional ODRL policy elements, such as prohibitions and duties, which are currently treated as fixed constraints. This extension requires more expressive reasoning mechanisms to capture complex dependencies and temporal aspects. In parallel, efforts will be directed toward integrating the framework into real-world data-sharing infrastructures, enabling automated, policy-compliant negotiation in practical systems such as data spaces and access control environments.

Acknowledgments

The work presented is based on research conducted within the Horizon Europe research and innovation programme CLIMRES (Grant Agreement No. 101147777). The content of the paper is the sole responsibility of its authors and does not necessarily reflect the views of the European Commission.

Declaration on Generative AI

The authors utilized ChatGPT-4 for paraphrasing and text condensation, and subsequently reviewed and revised the content, assuming full responsibility for the final version.

References

- [1] Mohammad Amini, Mohammad Fathian, and Mehdi Ghazanfari. A boa-based adaptive strategy with multi-party perspective for automated multilateral negotiations. *Applied Intelligence*, 50(9):2718–2748, Sep 2020.
- [2] Pallavi Bagga, Nicola Paoletti, Bedour Alrayes, and Kostas Stathis. A deep reinforcement learning approach to concurrent bilateral negotiation. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 297–303. International Joint Conferences on Artificial Intelligence Organization, 7 2020.
- [3] Debangshu Banerjee, Tarun Suresh, Shubham Ugare, Sasa Misailovic, and Gagandeep Singh. CRANE: Reasoning with constrained LLM generation. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu, editors, *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 2836–2857. PMLR, 13–19 Jul 2025.
- [4] Mate Bestek, Erik Grönvall, and Joanna Saad-Sulonen. Commoning semantic interoperability in healthcare. *International Journal of the Commons*, Aug 2022.
- [5] Zeyd Boukhers, Christoph Lange, and Oya Beyan. Enhancing data space semantic interoperability through machine learning: a visionary perspective. In *Companion Proceedings of the ACM Web Conference 2023, WWW ’23 Companion*, page 1462–1467, New York, NY, USA, 2023. Association for Computing Machinery.
- [6] Shengbo Chang and Katsuhide Fujita. Comb: Scalable concession-driven opponent models using bayesian learning for preference learning in bilateral multi-issue automated negotiation. *Group Decision and Negotiation*, 33(5):1143–1190, Oct 2024.
- [7] Andrea Cimmino, Juan Cano-Benito, and Raúl García-Castro. Open digital rights enforcement framework (odre): From descriptive to enforceable policies. *Computers & Security*, 150:104282, 2025.
- [8] Tobias Dam, Andreas Krimbacher, and Sebastian Neumaier. Policy patterns for usage control in data spaces, 2023.
- [9] Dave de Jonge. A new bargaining solution for finite offer spaces. *Applied Intelligence*, 53(23):28310–28332, Dec 2023.
- [10] Soulmaz Gheisari, Jaime Osvaldo Salas, Semih Yumusk, and George Konstantinidis. Minidm: Multi-issue negotiation in decentralised data marketplaces. In *Proceedings of the VLDB 2025 Workshops*, London, United Kingdom, 2025. VLDB Endowment.
- [11] Ryota Higa, Katsuhide Fujita, Toki Takahashi, Takumu Shimizu, and Shinji Nakadai. Reward-based negotiating agent strategies. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10):11569–11577, Jun. 2023.

- [12] Aaron Joyce and Vahid Javidroozi. Smart city development: Data sharing vs. data protection legislations. *Cities*, 148:104859, 2024.
- [13] George Konstantinidis, Adriane Chapman, Mark J. Weal, Ahmed Alzubaidi, Lisa M. Ballard, and Anneke M. Lucassen. The need for machine-processable agreements in health data management. *Algorithms*, 13(4), 2020.
- [14] Deuksin Kwon, Emily Weiss, Tara Kulshrestha, Kushal Chawla, Gale Lucas, and Jonathan Gratch. Are LLMs effective negotiators? systematic evaluation of the multifaceted capabilities of LLMs in negotiation dialogues. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 5391–5413, Miami, Florida, USA, November 2024. Association for Computational Linguistics.
- [15] Xudong Luo, Yanling Li, Qiaojuan Huang, and Jieyu Zhan. A survey of automated negotiation: Human factor, learning, and application. *Computer Science Review*, 54:100683, 2024.
- [16] Mashal Afzal Memon, Gian Luca Scoccia, and Marco Autili. A systematic mapping study on automated negotiation for autonomous intelligent systems. *Automated Software Engineering*, 32(2):44, May 2025.
- [17] Pankaj Mishra, Ahmed Maustafa, Takayuki Ito, and Minjie Zhang. Optimal auction based automated negotiation in realistic decentralised market environments. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(10):13726–13727, Apr. 2020.
- [18] Jihwan Oh. Llm agents for bargaining with utility-based feedback, 2025.
- [19] Irene Plaza-Ortiz, Andres Munoz-Arcentales, Joaquín Salvachúa, Carlos Aparicio, Gabriel Huecas, and Enrique Barra. Authentication and authorization in data spaces: A relationship-based access control approach for policy specification based on odrl, 2025.
- [20] Iyad Rahwan and Leila Amgoud. An argumentation based approach for practical reasoning. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '06, page 347–354, New York, NY, USA, 2006. Association for Computing Machinery.
- [21] Erwin J. Sacoto-Cabrera, Antonio Perez-Torres, Luis Tello-Oquendo, and Mariela Cerrada. Iot, ai, and digital twins in smart cities: A systematic review for a thematic mapping and research agenda. *Smart Cities*, 8(5), 2025.
- [22] Xiaofeng Wang, Zhixin Zhang, Jin Guang Zheng, Yiming Ai, and Rui Wang. Debt collection negotiations with large language models: An evaluation system and optimizing decision making with multi-agent. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Findings of the Association for Computational Linguistics: ACL 2025*, pages 9555–9577, Vienna, Austria, July 2025. Association for Computational Linguistics.
- [23] Ya Zhang, Jinghua Wu, and Ruiyang Cao. Optimizing automated negotiation: Integrating opponent modeling with reinforcement learning for strategy enhancement. *Mathematics*, 13(4), 2025.