

Data Sharing Negotiation and Contracting

Semih Yumusak^{1,*}, Soulmaz Gheisari^{1,*}, Jaime Osvaldo Salas¹, Syed Atif Moqurrab¹, Luis-Daniel Ibáñez¹ and George Konstantinidis¹

¹Department of Electronics and Computer Science, University of Southampton, Southampton, UK

Abstract

The domain of Data Marketplaces is an emerging technology and market field that monetises and trades data. The UPCASt project is developing a series of plugins for Data Marketplaces. The UPCASt Negotiation Plugin enhances data sharing efficiency by streamlining complex negotiations, and supporting and accelerating contract management between data providers and consumers. The plugin provides a centralised platform for initiating, managing, tracking, and finalising fine-grained negotiations on various conditions, while automating routine tasks and ensuring regulatory compliance. By utilising state-of-the-art negotiation protocols, such as the IDSA Contract Negotiation Protocol, and semantic technologies, such as ODRL and DPV, the system facilitates the creation of detailed, technology-independent contracts that are both machine-readable and human-readable, presenting a substantial advancement over traditional methods.

Keywords

data consumer, data contract, data marketplaces, data provider, data sharing agreements, negotiation

DEMO PAPER

1. Introduction

In the digital age, data has become a critical asset for driving innovation and enhancing decision-making. Organisations increasingly rely on data to optimise their operations and maintain a competitive edge. However, sharing and selling data over the web, or offline, introduces significant challenges related to compliance, security, and trust [1] [2]. Addressing these challenges requires well-structured agreements that safeguard all parties involved. Data sharing platforms can offer a comprehensive solution by facilitating these agreements and managing the complexities of data exchange. Through these platforms, organisations can protect their interests while leveraging shared data to gain insights, understand market trends, and foster collaborative advancements [3]. Thus, the role of data sharing platforms in enabling effective data exchange and collaboration is becoming increasingly vital.

The Semantic Web community has concentrated on defining data usage policies, for example through the ODRL W3C recommendation [4], which focuses on defining data usage and access control policies. However, interoperable enforcement across different implementations remains a significant challenge that the community continues to address. Moreover, data sharing

Posters, Demos, and Industry Tracks at ISWC 2024, November 13–15, 2024, Baltimore, USA

*Corresponding author.

✉ semih.yumusak@soton.ac.uk (S. Yumusak); s.gheisari@soton.ac.uk (S. Gheisari); j.o.salas@soton.ac.uk (J. O. Salas); s.a.moqurrab@soton.ac.uk (S. A. Moqurrab); l.d.ibanez@soton.ac.uk (L. Ibáñez); g.konstantinidis@soton.ac.uk (G. Konstantinidis)

🆔 0000-0002-8878-4991 (S. Yumusak); 0000-0001-8974-2841 (S. Gheisari); 0000-0002-9353-8955 (J. O. Salas); 0000-0002-8645-6687 (S. A. Moqurrab); 0000-0001-6993-0001 (L. Ibáñez); 0000-0002-3962-9303 (G. Konstantinidis)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

agreements between data providers and consumers encompass more than just terms of use or data usage policies; they may also require the definition and specification of the dataset to be shared [5], information on pricing, aspects of the environmental impact of datasets, and even details of specific data processing workflows (DPWs) that are going to process the data. To the best of our knowledge, these features have not been addressed by any data usage policy or negotiation framework. Furthermore, reaching an agreement requires effective and explicit management of the negotiation steps between data providers and data consumers, followed by both human-readable and machine-readable contract representations.

Achieving a successful negotiation on these features requires employing robust and strategic approaches. These approaches are essential to ensure fair and optimal outcomes for both data providers and consumers, balancing interests and aligning expectations effectively [6].

This paper presents the UPCASt system for negotiating and contracting data sharing agreements, which can work as a plugin in different data sharing platforms and different marketplaces. UPCASt facilitates data providers and consumers in negotiating dataset specifications, data processing workflows (DPW), usage policies (ODRL), and metadata like price and environmental impact. The system provides a centralized platform with APIs and web interfaces for managing negotiations and contracts, ensuring compliance with regulations like General Data Protection Regulation (GDPR)¹, and handling multi-party agreements.

Our tool extends the contract negotiation protocol defined by the International-Data-Spaces-Association (IDSA)², which in turn uses the ODRL concepts of Offer, Request and Agreement, which we extend, enabling more descriptive and technology-independent contracts. We use the ODRL ontology in combination with the Data Privacy Vocabulary (DPV)[7], which are used to define metadata as well as the use, processing and purpose of processing of data under relevant legislation, most notably GDPR. These vocabularies are utilised as standard primitives for our contract and negotiation language available to both sides of negotiation.

2. System Description and Demonstration

Before a negotiation starts we assume a provider has submitted their initial dataset specification and ODRL offers (specifying privacy constraints) in a data marketplace; these offers will be available for consumers to make requests upon.

Inline with A consumer initiates a negotiation by searching the available offers. All existing offers are accessible to consumers through their user interface in our plugin. Upon finding an offer for a resource, the consumer fetches the resource specification and its corresponding ODRL offer. The consumer then generates a request for negotiation, which consists of: (1) changing the resource specification of the resource, e.g., offering a lower price or different environmental requirements; (2) attaching a Data Processing Workflow (DPW) specification that describes the consumer's intended sequence of operations – any workflow tool can be used to generate the consumer's intended workflow here and we are currently implementing integration with nextFlow³ and a proprietary system; (3) changing the ODRL usage policy (initially attached on

¹<https://gdpr-info.eu/>

²<https://github.com/International-Data-Spaces-Association/ids-specification/blob/main/negotiation/contract.negotiation.protocol.md>

³<https://www.nextflow.io/>

the dataset) requesting changes on the purpose of usage, the intended actions, actors, duration of processing, etc. Note that the values of these policies are parameterisable: we currently use DPV and ODRL vocabularies to list these features but other ontologies, including domain specific ones (e.g., for negotiation of data sharing in a vertical data space such as manufacturing) is entirely possible; (4) lastly the consumer can attach a natural language “terms and conditions” part in their request. In our system either party can also set “negotiation preferences” which are the acceptable ranges within which they are willing to agree; these are not revealed to the other side and can be used to support automated negotiation by use of an external agent (we are currently developing one). In our system either party can also set “negotiation preferences” which are the acceptable ranges within which they are willing to agree; these are not revealed to the other side and can be used to support automated negotiation by use of an external agent (currently in development).

The provider may agree with the consumer’s request or propose a counteroffer, which, again, contains dataset specification details, a DPW pattern, ODRL usage policies, natural language part, and their negotiation preferences. A DPW pattern defines the preferred sequence for processing data from the provider’s side. Upon receipt of the counteroffer, the consumer continues with the same negotiation steps.

If the provider is the one that agrees to the consumer’s request at some step, the consumer is then notified and prompted to verify and sign, resulting in the establishment of an agreement. Note that both providers and the consumers can terminate the negotiation at any step. When the provider sends a counteroffer, and the consumer is the one accepting it⁴, then the provider is notified and, again, sends an agreement to be verified and signed by the consumer. Eventually, the verified agreement will be finalised and signed by the provider.

After the agreement is finalised, a contract is issued that includes the agreed final versions of the resource specification, DPW, ODRL usage policies and natural language terms and conditions. Metadata such as the start date and the validation period of the contract is also added and both parties sign it via their user interfaces. At the conference demo we will exhibit these steps on our web-interface which is build upon an API that implements the state changes in the negotiation process; this manages the entire negotiation process from initiation to termination or finalisation, while the provider/consumer interfaces offer user-friendly tools for initiating and managing negotiations, enable both parties to propose, modify, accept, or reject terms until they reach a mutual agreement. A video demonstrating the tool is available⁵.

2.1. Role of Semantic technologies

The UPCAST Negotiation Plugin uses semantic technologies like ODRL and DPV to automate and standardise the negotiation and contracting process. ODRL provides a flexible framework for specifying usage policies, essential for GDPR compliance, and streamlines the process by offering a common vocabulary, reducing inconsistencies and manual oversight. DPV further refines data privacy descriptions, enhancing the management of sensitive data sharing. Together, these technologies ensure clarity and consistency, minimizing conflicts and misunderstandings.

⁴Note that different terminology for the two sides such as *agree*, *accept*, or *verify* comes from the standardised IDSA protocol

⁵<https://github.com/EU-UPCAST/Demo/tree/main/Video>

Additionally we are currently implementing reasoning capabilities in our policy and negotiation engines.

3. Impact

We are currently implementing and evaluating the UPGAST Negotiation Plugin, using a real-world use case of a company processing and trading complex fitness data. Our partner company collects fitness data through devices and apps, offering insights into health, aiding personal and public health initiatives.

We are evaluating our technology using the following metrics:

- Agreement Completion Rate: Percentage of successful contract finalizations.
- User Satisfaction: Feedback on ease of use and overall experience.
- Regulatory Compliance: Adherence to data protection regulations (e.g., GDPR).
- Impact on Data-Sharing: Changes in data-sharing behavior and agreement fairness.

At the same time, we are evaluating the following issues regarding users' experience:

User Interface Considerations: The plugin's user-friendly web interface guides users through negotiation with minimal technical expertise.

Potential Feedback Collection: We'll gather feedback during pilot tests to refine the interface, focusing on navigation, clarity, and overall satisfaction.

Impact on Data-Sharing Practices: Simplifying negotiations and ensuring compliance, the plugin aims to foster a collaborative environment, encouraging broader data-sharing and improving decision-making across sectors.

4. Conclusion

The UPGAST Negotiation and Contracting system offers a promising approach to managing data-sharing agreements by integrating state-of-the-art negotiation protocols with semantic technologies, using generic or domain-specific vocabularies to streamline the negotiation, agreement and contracting process and present human and machine-readable automated contracts. By demonstrating how semantic technologies like ODRL and DPV can be applied in practical data-sharing scenarios, this work contributes valuable insights into the future development of data marketplaces and the broader Semantic Web ecosystem.

Ongoing improvements include advanced reasoning engines for conflict resolution, domain-specific ontologies for specialized sectors, and semantic matching algorithms to optimise negotiation strategies based on past agreements.

Acknowledgments

This work was funded by the UKRI Horizon Europe guarantee funding scheme for the Horizon Europe projects UPGAST (101093216101093216) and RAISE (101093216101058479).

References

- [1] C. Caimi, C. Gambardella, M. Manea, M. Petrocchi, D. Stella, Legal and technical perspectives in data sharing agreements definition, in: Annual Privacy Forum, Springer, 2015, pp. 178–192.
- [2] J. Byabazaire, G. O’Hare, D. Delaney, Data quality and trust: Review of challenges and opportunities for data sharing in iot, *Electronics* 9 (2020). URL: <https://www.mdpi.com/2079-9292/9/12/2083>. doi:10.3390/electronics9122083.
- [3] I. Jussen, J. Schweihoff, V. Dahms, F. Möller, B. Otto, Data sharing fundamentals: characteristics and definition, in: Proceedings of the 56th Hawaii International Conference on System Sciences, 2023.
- [4] R. Ianella, Open digital rights language (odrl), *Open Content Licensing: Cultivating the Creative Commons* (2007).
- [5] M. Li, Dsap: Data sharing agreement privacy ontology, Ph.D. thesis, 2018.
- [6] S. Fatima, S. Kraus, M. Wooldridge, *Principles of automated negotiation*, Cambridge University Press, 2014.
- [7] H. J. Pandit, B. Esteves, G. P. Krog, P. Ryan, D. Golpayegani, J. Flake, Data privacy vocabulary (dpv)–version 2, arXiv preprint arXiv:2404.13426 (2024).