Semantic Web 15 (2024) 643–645
DOI 10.3233/SW-243615
IOS Press

Editorial of the special issue on semantic technologies for data and algorithmic governance

Sabrina Kirrane[a,*], Oshani Seneviratne[b] and Michel Dumontier[c]
[a] *Vienna University of Economics and Business, Austria*
E-mail: sabrina.kirrane@wu.ac.at
[b] *Rensselaer Polytechnic Institute, USA*
E-mail: senevo@rpi.edu
[c] *Maastricht University, The Netherlands*
E-mail: michel.dumontier@maastrichtuniversity.nl

Keywords: GDPR, privacy, trust, transparency, compliance, accountability

1. Preface

Technology is playing a progressively important key role in enabling effective governance structures, processes, and frameworks. As society increasingly depends on complex systems ranging from simple ‘decision support systems’ to ‘systems of systems’ and ‘semi-autonomous systems’, data and algorithmic governance are paramount. When it comes to data and algorithmic governance tools and techniques, there are still many open questions. For instance, to what extent do these systems safeguard against privacy violations and honour intellectual property rights? Can granular consent be granted, and will its users understand the consequences? How can biases, discrimination, and censorship be identified and acted on? Do transparency and explainability lead to greater accountability? How can trust be woven into the fabric of these systems? Towards this end, this special issue aims to explore the development and evaluation of semantic technologies with respect to the data and algorithmic governance mechanisms, processes, and methodologies that are critically needed to support the development of trust-centric social and business applications.

2. Topics of interest

The special issue solicited high-quality submissions on (but are not restricted to) the following topics:

- Findable, Accessible, Interoperable, and Reusable (FAIR) data management
- Techniques for enabling ownership, control, and access
- Identifying fake news and misinformation
- Managing bias and ensuring fairness
- Enabling transparency, explainability, and accountability

[*] Corresponding author. E-mail: sabrina.kirrane@wu.ac.at.

- Methods for policy governance
- Information flow control and accountability
- Measuring data quality
- Managing the data life cycle
- Metrics for assessing the effectiveness of governance algorithms
- Data privacy, regulations, and compliance
- Provenance, trust, and metadata for authoritative sources
- Privacy and security enforcement
- Methods for information flow control and accountability
- Frameworks and systems for personal data storage and control
- Ensuring data authenticity and integrity
- Privacy-preserving data mining and machine learning methods
- Protecting against identity theft and data falsification
- User-friendly interface design for data and algorithmic governance
- Standards for data and algorithmic governance
- Tackling legal issues with respect to data and algorithms
- Law and governance in e-democracy and e-participation
- Benchmarking approaches to data and algorithmic governance
- Building trust and transparency mechanisms in the fabric of the Web
- Participatory frameworks for fair and efficient algorithmic governance

3. Content

In the following, we provide a high-level overview of the papers comprising this special issue:

Consent Through the Lens of Semantics: State of the Art Survey and Best Practices [4] by *Anelia Kurteva, Tek Raj Chhetri, Harshvardhan J. Pandit, and Anna Fensel* is a systematic literature review of existing papers, projects, and standardisation efforts that use semantic technology to implement consent based on requirements stipulated in the European Union (EU) General Data Protection Regulation (GDPR).[1] The authors propose a consent life-cycle that incorporates five different states: the *request* for consent, the *comprehension* by the data subject with respect to what they are consenting to, the *decision* to give or refuse consent, and the ongoing *use* of personal data and the *consent management*. The life cycle is subsequently used to identify best practices for industry, researchers, and policymakers. Based on the analysis performed, the authors conclude that a semantic model for consent is highly beneficial as it allows for a common understanding of consent requirements that both humans and machines can understand, facilitates transparency and risk assessment, and caters to ongoing compliance checking and consent management.

Analysis of Ontologies and Policy Languages to Represent Information Flows in GDPR [3] by *Beatriz Esteves and Víctor Rodríguez-Doncel* also investigates how semantic technology can be used to satisfy GDPR requirements, however the focus is on the use of ontologies and policy languages. Guided by information flows derived from the rights and obligations stipulated in the GDPR, the authors perform an integrated literature review with the goal of assessing the suitability of existing policy languages and ontologies for modeling GDPR requirements with respect to information flows. The authors conclude that LegalRuleML ,[2] the Open Digital Rights Language (ODRL),[3] the Data Privacy Vocabulary (DPV),[4] and GDPR text extensions (GDPRtEXT)[5] have reached a level of maturity that makes them suitable for representing GDPR rights and obligations. These resources were highlighted

[1] GDPR, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679&qid=1681238509224.

[2] LegalRuleML, https://docs.oasis-open.org/legalruleml/legalruleml-core-spec/v1.0/legalruleml-core-spec-v1.0.pdf.

[3] ODRL, https://www.w3.org/TR/odrl-model/.

[4] DPV, https://w3c.github.io/dpv/dpv/.

[5] GDPRtEXT, https://www.w3.org/community/dpvcg/wiki/GDPRtEXT.

as together they can be used to represent many of the GDPR information flows, are openly available, and are rooted in standardisation efforts.

Semantic-enabled Architecture for Auditable Privacy-Preserving Data Analysis [2] by *Fajar J. Ekaputra, Andreas Ekelhart, Rudolf Mayer, Tomasz Miksa, Tanja Šarčević, Sotirios Tsepelakis, and Laura Waltersdorfer* proposes a semantic enables architecture, entitled WellFort, which facilitates secure storage of consent and allows for the execution of privacy-preserving data analytics processes. The effectiveness of the proposed conceptual architecture is demonstrated via a semantic technology-based instantiation and the subsequent evaluation, which was conducted with the help of four separate healthcare use cases. The authors highlight that technological approaches for automated usage compliance checking are already quite mature compared to other features. Future work includes the adaption of the prototype to cater to additional use cases and the development of a multi-purpose audit toolbox to support privacy-preserving data analytics over heterogeneous sources.

Differential Privacy and SPARQL [1] by *Carlos Buil Aranda, Jorge Lobo, and Federico Olmedo* investigates if differential privacy techniques designed for relational databases and SQL joins can be applied to SPARQL counting queries over Resource Description Framework (RDF) knowledge graphs. The proposed algorithm is realised via a differential privacy query engine that uses approximation in order to answer SPARQL counting and grouping queries. The engine's potential is demonstrated via various simulations derived from real-world data and queries from Wikidata. Future work involves supporting additional operations (e.g., sums and averages) and analysing the impact differential privacy algorithms have on SPARQL query engines.

Interestingly, all papers were influenced (to a lesser or greater extent) by the GDPR. It is worth noting that the proposed consent mechanisms, ontologies, policy languages, and privacy-preserving technologies could also be used to satisfy some requirements stipulated in the proposed EU Data Governance Act,[6] which puts a major focus on consent, and the EU Artificial Intelligence (AI) act,[7] which stipulates restrictions with regard to the processing of personal data. Unfortunately, the coverage of the topics of interest for this special issue was limited, suggesting that data and algorithm governance research needs greater attention.

References

[1] C. Buil-Aranda, J. Lobo and F. Olmedo, Differential privacy and SPARQL, *Semantic Web* (2021), 1–34.

[2] F.J. Ekaputra, A. Ekelhart, R. Mayer, T. Miksa, T. Šarčević, S. Tsepelakis and L. Waltersdorfer, Semantic-enabled architecture for auditable privacy-preserving data analysis, *Semantic Web* (2021), 1–34.

[3] B. Esteves and V. Rodríguez-Doncel, Analysis of ontologies and policy languages to represent information flows in GDPR, *Semantic Web* (2022), 1–35.

[4] A. Kurteva, T.R. Chhetri, H.J. Pandit and A. Fensel, Consent through the lens of semantics: State of the art survey and best practices, *Semantic Web* (2021), 1–27. doi:10.3233/SW-210438.

[6] Governance Act, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.

[7] AI Act, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.