Legal Knowledge and Information Systems
J. Savelka et al. (Eds.)

doi:10.3233/FAIA241258

Ontology-Based Approach for Mapping Concepts and Requirements from Regulations and Standards: The Case of the EU AI Act and International Standards

Julio HERNANDEZ [a,1], Delaram GOLPAYEGANI [a] and David LEWIS [a]
[a] *ADAPT Centre, School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland.*

Abstract. The many initiatives on trustworthy AI result in a confusing and multipolar landscape that organizations are operating within the fluid and complex international value chains must navigate in pursuing trustworthy AI. The EU's proposed AI Act will now shift the focus of these organizations towards the normative requirements for regulatory compliance. Understanding to what extent standards compliance will deliver regulatory compliance for AI remains a complex challenge. This paper introduces the Trustworthy AI Requirements (TAIR) ontology, a simple and replicable method for extracting and sharing relevant terms and concepts from legal regulations and standard texts into open-knowledge graphs. The TAIR ontology is vital for evaluating the sufficiency of standards conformance for regulatory compliance, providing a foundation for identifying areas where further development of technical consensus in trustworthy AI value chains will be indispensable to achieve regulatory compliance. The evaluation of the TAIR ontology aims to detect errors or inconsistencies based on best practices for ontology design.

Keywords. Open Knowledge Graph, Trustworthy AI, AI Act, Standards, Legal Compliance

1. Introduction

The global interest in the ethical and social risks of AI has grown rapidly in recent years [1,2,3,4]. In the primary wave of trustworthy AI initiatives, guidelines typically are presented as structured statements of principles that organizations can adopt to demonstrate some degree of trustworthiness in their development and use of AI technology[5]. With the increasing number of AI incidents[6,7,8], it became evident for policymakers and public authorities that there is a wide range of applications through which AI negatively impacts people's lives that are developed and deployed with little external oversight [9,10].

With its political agreement on the AI Act [11] being reached at the end of 2023, the European Union (EU) has become a pioneer in AI regulation. The AI Act specifies

[1] Corresponding Author: Julio Hernandez, julio.hernandez@adaptcentre.ie.

a tiered risk system, where some applications of AI are *prohibited* (Unacceptable risk), and others are identified as a sufficiently low risk that only consumer labels or voluntary codes of practice are required (Limited and Minimal risk). However, the focus of the Act relies on regulatory oversight and compliance information exchange between these tiers where high-risk AI systems are defined.

With the forthcoming enforcement of the AI Act, one of the key challenges for high-risk AI providers and deployers is to navigate a sea of standards to address trustworthy AI requirements through regulatory compliance. Additionally, the lack of common terminology and detailed mapping of requirements are added to the complexity faced by providers and deployers. Any mappings between legal requirements for trustworthy AI and technical standards that enable conformance and certification functions that satisfy those requirements should be flexible, extensible, transparent, and auditable solutions to satisfy regulatory and organizational rules on governance process integrity. Open standards should be used, as far as possible, to increase third-party inspection and, therefore, confidence in the completeness and accuracy of mapping. In this paper, we take an approach based on Open Knowledge Graphs (OKG) specified using standards from the World Wide Web Consortium (W3C), which have been proven to be successful in promoting interoperability between approaches, satisfying the requirements of the EU General Data Protection Regulation (GDPR) [12] and expressing high-risk information through an AI risk ontology based on the requirements of the AI Act and the ISO 31000 series of standards [13].

2. Related work

Some existing work addresses the challenges of implementing trustworthy AI requirements using OKG-based approaches. Amaral et al. [14] combine the Reference Ontology for Trust (ROT) and the Non-Functional Requirements Ontology (NFRO) to characterize an ontology that captures trust requirements for software systems. Inspired by ISO/IEC JTC 1/SC 42 activities, Lewis et al. [15] propose a high-level ontology to map out the consistency and overlap of concepts from different AI standards, regulations, and policies. Golpayegani et al. [16] use the aforementioned ontology to compare the semantic interoperability between ISO/IEC 42001 standard on AI management system, the EU trustworthy AI assessment list (ALTAI) and the EU AI Act. In this work, are map AI concepts and requirements from regulations and standards to develop a mechanism to compare, integrate, and relate the terminology used by these documents with the objective of regulatory compliance.

3. TAIR: Trustworthy AI Requirements Ontology

The Trustworthy AI Requirements (TAIR) ontology[2] provides the elements to describe concepts and requirements associated with a regulation or standard. **Figure 1** depicts the TAIR ontology, where `Requirement` and `Concept` are the main classes in the ontology.

[2]TAIR webpage: https://tair.adaptcentre.ie/

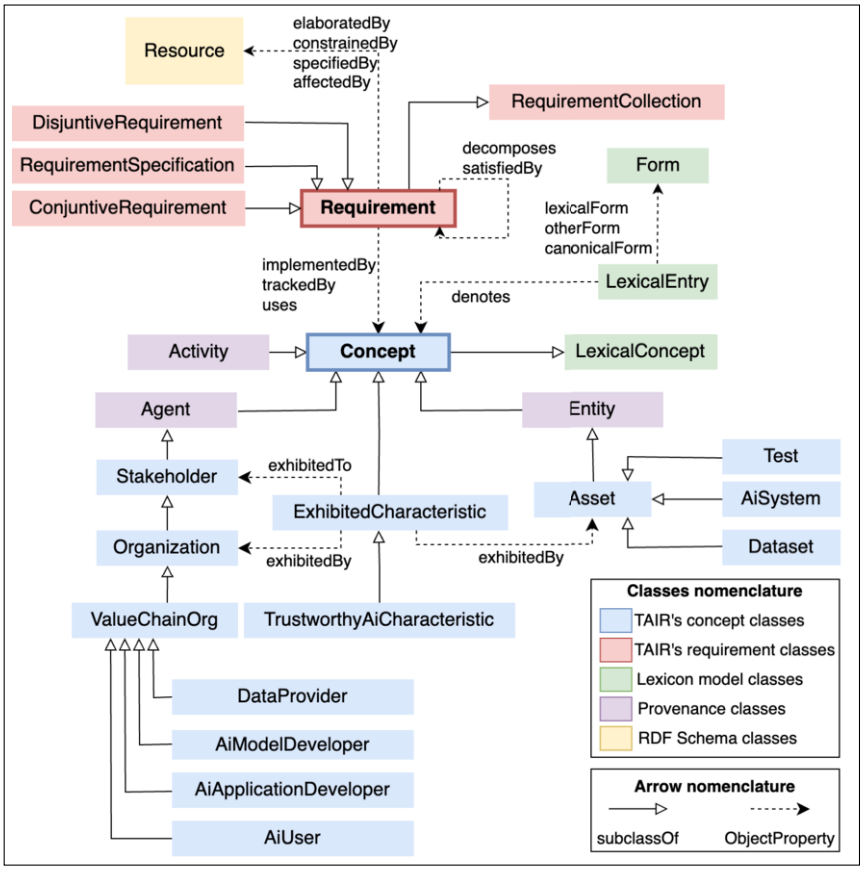

Figure 1. The TAIR ontology - key concepts and relations.

The `Concept` class is a subclass of the `OntoLex`[3] vocabulary, which describes linguistic resources such as the representation of dictionaries or annotations commonly found in lexicography. The `Requirement` class is used to describe normative clauses. A requirement could be related to a particular concept or lexical entry; this relationship is denoted by the properties `implementedBy` (who is responsible for implementing the described requirement), `trackedBy` (who tracks the updates of the requirement), and `uses` (who uses the described requirement).

3.1. Requirements and Concepts Semantic Mappings

The mapping process (Figure 2) considers the regulation or standard document structure divided into clauses. The following paragraphs describe the three phases of semantic mapping.

Elements identification In this phase, the concepts and requirements from a regulation or standard are identified. Concepts are usually defined in a special section called "Terms

[3] `https://www.w3.org/2019/09/lexicog/`

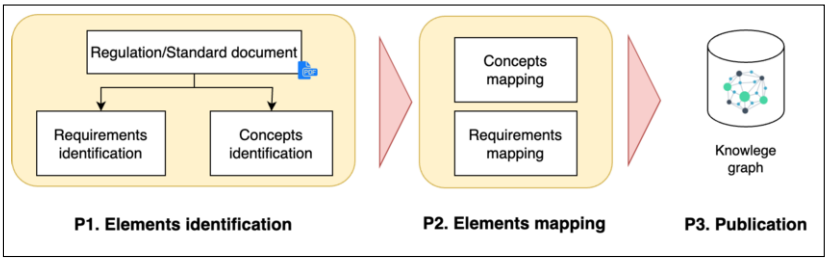

Figure 2. The three phases in the regulation and standard requirements and concepts mapping process.

and definitions" or "Definitions". On the other hand, requirements identification consists of looking for clauses expressed in the verbal form of shall or shall not [4]

Elements mapping After identifying the concepts and requirements, the next phase describes each requirement and concept definition into a linked data element, considering the classes and properties from the TAIR ontology. The mapping process is divided into concepts, lexical entries, and requirements, which are explained in the following paragraphs.

i) Concepts mapping. The concept extraction and mapping process first involves extracting explicitly defined terms such as SKOS concepts. The Simple Knowledge Organization System (SKOS) [17] can organize concepts into concept sets and establish hierarchical relationships useful for building taxonomies. In SKOS, hierarchical associations are defined as a 'narrower' or 'broader' relationship between concepts. The structure of terminological lists (for example, subsection in the terminology section of ISO/IEC standards), the text of the definitions, and cross-references between these are used to capture taxonomical structures, using the SKOS 'narrower', 'broader', and 'related' relationships.

ii) Lexical entries mapping. Lexical entries are candidates for alignment with definitions from another document, e.g., from another referenced legislative document or technical standard.

iii) Requirements mapping. Normative clauses are converted to atomic normative requirements[5].

iv) Requirement and concept mapping. A concept is associated with a specific requirement through the properties `uses` if it is directly mentioned in the requirement.

Publication This phase provides the mechanisms for accessing the ontology documentation and querying the requirements and concepts. The Ontotext GraphDB[6] graph database was used to publish the TAIR ontology. GraphDB is a triplestore with RDF and SPARQL support and graph visualization capabilities. Two demos [7] of the TAIR ontology were developed, focusing on the requirements and concepts of the AI Act. The first demo explores Chapter III of the AI Act related to High-Risk AI System requirements. The second demo explores the concepts defined by the AI Act.

[4] ISO/IEC Directives, Part 2 - https://www.iso.org/sites/directives/current/part2/index.xhtml

[5] A specific irreducible requirement involving named actors, activities, or entities

[6] https://graphdb.ontotext.com/

[7] TAIRdemo:https://tair.adaptcentre.ie/demo.html

The extraction of requirements from the AI Act related to compliance obligations on AI providers resulted in 118 separate requirements. Where relevant, these are linked to the 46 explicitly defined concepts from Article three of the AI Act. Additionally, 23 lexical entries were extracted from the AI Act requirements.

3.2. Ontology evaluation

The TAIR ontology language conformity evaluation was conducted through the OntOlogy Pitfall Scanner! (OOPS!) tool [18]. The OOPS! tool detects potential problems in the provided ontology by means of a semiautomatic diagnosis for 32 pitfalls. Based on the detected pitfall, the evaluation result is classified as minor, important, and critical. Each pitfall is associated with an importance level decided in conjunction with OOPS! developers, experienced ontological engineers, and users.

The pitfalls identified by the OOPS! tool for the TAIR ontology are minor problems. The most recurrent pitfall is the missing definition of inverse relationships, e.g., the inverse property `constrains` is not defined for the property `constrainedBy`. The missing annotation pitfalls refer to properties and/or classes without a human-readable property; it mainly occurs for external classes defined in the TAIR ontology, such as `LexicalConcept` or `Resource` classes. Finally, the unconnected ontology elements pitfall occurs because a defined class is not connected with any other element of the ontology, e.g., the class `LexicalConcept` is not connected with any other class; the class `Concept` refers to it but only as their subclass. All the unconnected ontology elements and missing annotation pitfalls reference external vocabularies, e.g., SKOS or RDFS; their definition will be found in the corresponding URI.

4. Conclusion and Future Work

The Trustworthy AI Requirements (TAIR) ontology provides a basis for mapping concepts and requirements from normative statements in the AI Act and the conformance-focused international standard on AI from SC42. It is partially available as an Open Knowledge Graphs (OKG) resource that relates relevant concepts and requirements published in a traceable, queryable, and navigable manner.

The model may be of use to policymakers and standards developers involved in the development of harmonized standards, in guidelines to support the implementation of the Act, such as EC guidelines to SME developing or public sector agencies procuring AI, and those establishing transparency mechanisms for regulatory learning mechanisms such as regulatory sandboxes and real-life trials.

In the long term, this approach and its open resources could be used to compare proprietary or national trustworthy AI mechanisms to the conformance and compliance system offered by the AI Act and its harmonized standards.

Acknowledgement

This project has received funding as a research gift from Meta and is supported by the Science Foundation Ireland under Grant Agreement No 13/RC/2106_P2 at the ADAPT SFI Research Centre and the European Union's Horizon 2020 Marie Skłodowska-Curie grant agreement No 813497 for the PROTECT ITN.

References

[1] Yigitcanlar T, Desouza KC, Butler L, Roozkhosh F. Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature. Energies. 2020;13(6):1473.

[2] Nasim SF, Ali MR, Kulsoom U. Artificial intelligence incidents & ethics a narrative review. International Journal of Technology, Innovation and Management (IJTIM). 2022;2(2):52-64.

[3] Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, et al. An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Ethics, governance, and policies in artificial intelligence. 2021:19-39.

[4] O'Reilly-Shah VN, Gentry KR, Walters AM, Zivot J, Anderson CT, Tighe PJ. Bias and ethical considerations in machine learning and the automation of perioperative risk assessment. British journal of anaesthesia. 2020;125(6):843-6.

[5] Li B, Qi P, Liu B, Di S, Liu J, Pei J, et al. Trustworthy AI: From principles to practices. ACM Computing Surveys. 2023;55(9):1-46.

[6] Lee NT, Resnick P, Barton G. Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. 2019.

[7] Anderson JM, Nidhi K, Stanley KD, Sorensen P, Samaras C, Oluwatola OA. Autonomous vehicle technology: A guide for policymakers. Rand Corporation; 2014.

[8] Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. ACM computing surveys (CSUR). 2021;54(6):1-35.

[9] Beckman L, Hultin Rosenberg J, Jebari K. Artificial intelligence and democratic legitimacy. The problem of publicity in public authority. AI & SOCIETY. 2022:1-10.

[10] The European Commission. COMMUNICATION FROM THE COMMISSION. The European Commission; 2018. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN.

[11] The European Commission. REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. The European Commission; 2021. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.

[12] Pandit HJ, O'Sullivan D, Lewis D. Queryable Provenance Metadata For GDPR Compliance. Procedia Computer Science. 2018;137:262-8. Proceedings of the 14th International Conference on Semantic Systems 10th – 13th of September 2018 Vienna, Austria. Available from: https://www.sciencedirect.com/science/article/pii/S1877050918316314.

[13] Golpayegani D, Pandit HJ, Lewis D. AIRO: An Ontology for Representing AI Risks Based on the Proposed EU AI Act and ISO Risk Management Standards. In: International Conference on Semantic Systems; 2022. Available from: https://api.semanticscholar.org/CorpusID:252919566.

[14] Amaral G, Guizzardi R, Guizzardi G, Mylopoulos J. Ontology-based modeling and analysis of trustworthiness requirements: Preliminary results. In: International Conference on Conceptual Modeling. Springer; 2020. p. 342-52.

[15] Lewis D, Filip D, Pandit HJ. An Ontology for Standardising Trustworthy AI. In: Hessami AG, Shaw P, editors. Factoring Ethics in Technology, Policy Making, Regulation and AI. Rijeka: IntechOpen; 2021. Available from: https://doi.org/10.5772/intechopen.97478.

[16] Golpayegani D, Pandit HJ, Lewis D. Comparison and Analysis of 3 Key AI Documents: EU's Proposed AI Act, Assessment List for Trustworthy AI (ALTAI), and ISO/IEC 42001 AI Management System. In: Irish Conference on Artificial Intelligence and Cognitive Science. Springer; 2022. p. 189-200.

[17] Isaac A, Summers E. SKOS simple knowledge organization system primer. Working Group Note, W3C. 2009.

[18] Poveda-Villalón M, Gómez-Pérez A, Suárez-Figueroa MC. OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation. International Journal on Semantic Web and Information Systems (IJSWIS). 2014;10(2):7-34.