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*

Delaram Golpayegani[0000−0002−1208−186X], Harshvardhan J. Pandit[0000−0002−5068−3714], and Dave Lewis[0000−0002−3503−4644]

ADAPT Centre, School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland {sgolpays, pandith, delewis}@tcd.ie

Abstract. Conforming to multiple and sometimes conflicting guidelines, standards, and legislations regarding development, deployment, and governance of AI is a serious challenge for organisations. While the AI standards and regulations are both in early stages of development, it is prudent to avoid a highly-fragmented landscape and market confusion by finding out the gaps and resolving the potential conflicts. This paper provides an initial comparison of ISO/IEC 42001 AI management system standard with the EU trustworthy AI assessment list (ALTAI) and the proposed AI Act using an upper-level ontology for semantic interoperability between trustworthy AI documents with a focus on activities. The comparison is provided as an RDF resource graph to enable further enhancement and reuse in an extensible and interoperable manner.

Keywords: Trustworthy AI · AI management system · ALTAI · AI Act · ISO/IEC 42001 · Ontology · Activity · Comparison.

1 Introduction

The wide application of AI systems urges governments, legislators, standardisation bodies, and think tanks to encourage and sometimes obligate organisations to develop and use AI in a trustworthy manner. AI regulations, standards, and guidelines developed separately and in isolation risk a highly fragmented landscape that can lead to regulatory and market confusion. Consequently, organisations are compelled to navigate a large number of competing and changing

* This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813497, as part of the ADAPT SFI Centre for Digital Media Technology is funded by Science Foundation Ireland through the SFI Research Centres Programme and is co-funded under the European Regional Development Fund (ERDF) through Grant#13/RC/2106_P2. Harshvardhan J. Pandit has received funding under the Irish Research Council Government of Ireland Postdoctoral Fellowship Grant#GOIPD/2020/790.

requirements from multiple sources regarding AI development and use. The lack of alignment between different sources of requirements, such as laws and standards, creates difficulties in identifying and fulfilling obligations.

In this paper, we identify the commonality, inconsistencies, and gaps across the following three dominant AI documents within the scope of EU's regulatory regime: the proposed AI Act [1], Assessment List for Trustworthy AI (ALTAI) [2], and the draft ISO/IEC 42001 standard for AI management systems[1].

Amongst these three, we utilise ISO/IEC 42001 as the primary source of requirements given its distinct role as a certifiable standard, and compare the others with it to indicate adherence towards guidelines (ALTAI) and regulations (AI Act). More specifically, we investigate the following questions:

(i) To what extent can ALTAI's trustworthy AI requirements be integrated into ISO/IEC 42001's AI management system activities?
(ii) To what extent can AI Act's high-risk AI obligations be integrated into ISO/IEC 42001's AI management system activities?

We address the aforementioned questions by proposing a methodology to compare AI documents using an upper-level trustworthy AI ontology [3], which enables modelling and linking concepts within AI documents (see Section 2). We then demonstrate the comparison of ISO/IEC 42001 with ALTAI's trustworthy AI (Section 3) and the AI Act (Section 4). The comparison is made available online as an RDF resource to enable further enhancement and reuse[2]. We discuss semantic modelling of activities extracted from the documents in Section 5. In Section 6, related work on ontology-based comparison of policies, regulations, and standards is mentioned and we conclude the paper and identify avenues for future work in Section 7.

2 Methodology for Comparison and Analysis

AI documents can be compared on the basis of different semantic building blocks: key terms defined within them, activities mentioned, and normative requirements or obligations required to be met for compliance. Considering the central focus of management system standards on organisational activities and processes, we limit the scope of our comparison to activities.

Given that different standards, regulations, and policies are being created for evaluating trustworthiness of AI, there is bound to be some overlap between them. To assist in the task of comparing them, a conceptual model and framework is essential to identify and link together the relevant concepts within different documents. An ontological representation permits formalisation of the conceptual model and its application in use-cases. With this view, Fig. 1 presents the core ontology for supporting mapping of concepts between different emerging AI standards. It is based on activities carried out within ISO/IEC (more

[1] https://www.iso.org/standard/81230.html
[2] https://github.com/delaramglp/aidocs

specifically sub-committee 42) regarding AI standardisation and incorporates existing ISO/IEC standards and outputs for 'characteristics' expressed by trustworthy AI systems. The premise of the ontology rests on the fact that several

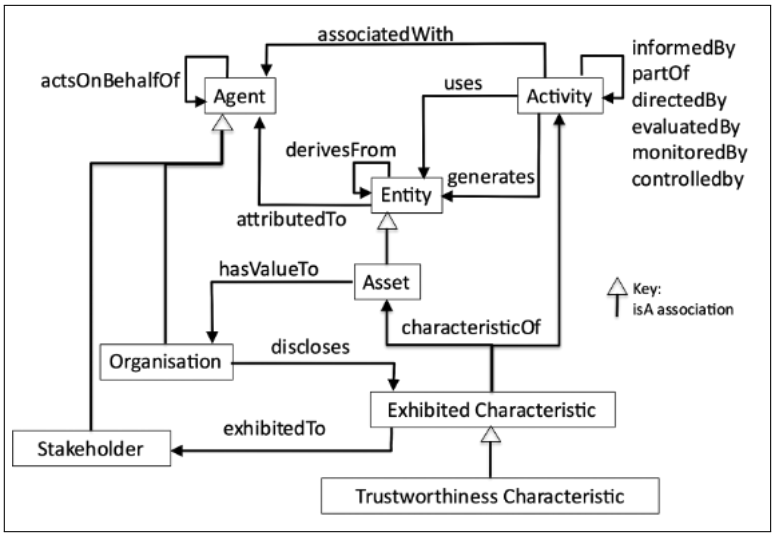

Fig. 1. Upper-level ontology for mapping trustworthy AI documents [3]

of trustworthy characteristics are yet to be clarified and defined in relation to AI and AI development activities. Therefore, it focuses on specifying the relationships between activities, entities, assets, and characteristics (exhibited for trustworthiness), agents, stakeholders, and organisations. The ontology is based on Basic Formal Ontology (BFO) - a generic upper-level ontology used in formalisations across domains, and the PROV-O ontology which is a W3C standard for expressing provenance.

The ontology provides a way to express activities of organisations that relate to AI where the trustworthiness is manifested through characteristics of Entities that make up a product or service employing AI. It also provides a way to depict the influence of entities, activities, and agents in these processes, and captures the role of stakeholders in disclosing and exhibiting trustworthiness of AI through its characteristics. The ontology thus enables representing use of AI from both within and outside the perspective of an organisation or service, and is useful for comparing different AI guidelines by using its conceptual model as a framework for identifying and aligning concepts.

We utilise the trustworthy AI ontology to compare AI documents in order to assess the degree of alignment between them by modelling and linking trustworthy AI activities mentioned within them. The following describes the steps taken for analysis and comparison of documents:

1. The documents are analysed to extract relevant activities to trustworthy AI, which then modelled as `Activity`.
2. `partOf` relationship is used to bridge the isolated sets of `Activities` identified from the documents.
3. An analysis is carried out to identify the overlaps and potential conflicts through investigation of activities that are mapped or could not be mapped using the `partOf` relation.

3 Comparison of ALTAI with ISO/IEC 42001

3.1 ALTAI Activities

ALTAI suggests a set of questions, grouped by the ethical principle under assessment, for assessing whether an AI system adheres to trustworthy AI requirements specified in [4] (see the structure of ALTAI in Fig. 2). Designed for trustworthy AI self-assessment, ALTAI provides useful hints regarding development and use of AI systems. One of the aspects of trustworthiness assessment is execution of particular activities; for example, 'Are end-users or other subjects adequately made aware that a decision, content, advice or outcome is the result of an algorithmic decision?', which is a question listed under Human Agency and Oversight requirements, implies execution of an activity to *inform end-users or other subjects that a decision, content, advice or outcome is the result of an algorithmic decision*. For the purpose of comparison, we made the management activities implied by ALTAI questions explicit.

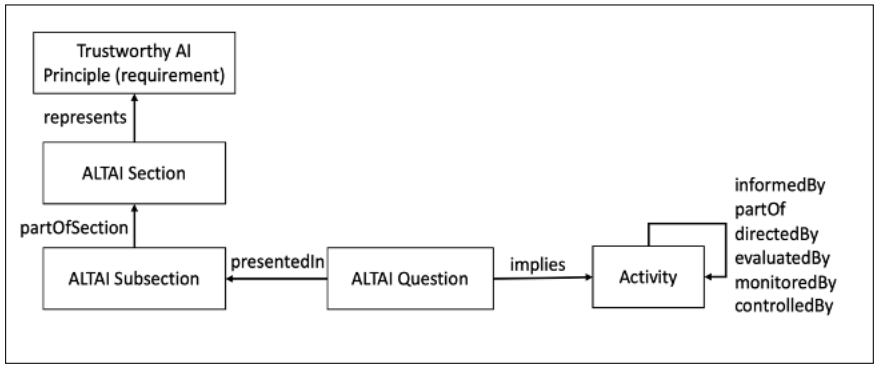

Fig. 2. ALTAI structure

3.2 AI Management System Activities

The ISO/IEC 42001 standard for AI management systems, being developed by JTC 1/SC 42, is currently (Nov'22) in DIS or draft stage, implying relative maturity awaiting final comments before publication. It follows the 'harmonised

structure' of all management system standards developed by ISO, which is defined in the openly available ISO/IEC Directives on procedures for ISO technical work[3]. Based on the harmonised structure, Lewis et al. [3] identified AI management system activities, where each is given an identifier, a label, and a 'see also' attribute which is a link to the relevant harmonised structure clause. The entities generated and used by each activity are represented in a similar manner. The updated list of AI management system activities, which reflects the latest version of the Directive published in 2022, is presented in Table 1.

Table 1. AI Management System (AIMS) activities

o No.	ID	AIMS activity (label)	HS clause (see also)
1	UOC	Understanding organisation and its context	4.1
2	USE	Understanding stakeholder needs and expectation	4.2
3	DS	Determine AIMS scope	4.3
4	EIMI	Establish, implement, maintain and continually improve management system and its processes	4.4
5	DLC	Demonstrate leadership and commitment to the management system	5.1
6	EP	Establish AIMS policy	5.2
7	ARRA	Assign roles, responsibilities and authorities	5.3
8	ARO	Address risks and opportunities	6.1
9	EPAO	Establish and plan to achieve AI objectives	6.2
10	ARRA	Assign roles, responsibilities and authorities	6.3
11	DAR	Determine and allocate resources for AIMS	7.1
12	DEC	Determine and ensure competence of people affecting AI performance	7.2
13	PA	Promote awareness	7.3
14	DC	Determine AIMS communication	7.4
15	CUCD	Create, update, and control documented information	7.5
16	PCP	Plan and control AI processes	8.1
17	MMAE	Monitor, measure, analyse and evaluate AI	9.1
18	IA	Internal (AIMS) audit	9.2
19	UMR	Undertake management review	9.3
20	DNCA	Detect non-conformance and take corrective action	10.1
21	CI	AIMS Continual improvement	10.2

3.3 ALTAI - ISO/IEC 42001 Activity Comparison

By comparing ALTAI with ISO/IEC 42001, we aim to investigate the following:

– Is there any organisational activity required for trustworthy AI that cannot be integrated into an AI management system?

[3] `https://www.iso.org/sites/directives/current/consolidated/index.xhtml`

- Which AI management systems activities do not play a role in achieving trustworthiness?
- What management systems activities are involved in achieving a particular trustworthy AI requirement, e.g. privacy and data governance?

Alignment Groups In the comparison process, a number of commonly occurring structures are identified. For instance, multiple ALTAI activities that refer to achieving AI objectives such as *Accuracy*, *Explainability*, *Privacy*, and *Fairness* are `partOf` 'establish and plan to achieve AI objectives' activity. We categorise these structures into the 17 alignment groups listed in Table 2.

Table 2. ALTAI - AI management system activities alignment groups

ID	ALTAI activity structure	partOf (AIMS activity)
AG1	Assess the impact of the AI system	ARO
AG2	Assess the system vulnerabilities or threats	ARO
AG3	Assess whether the AI system respects a specific right	ARO
AG4	Establish processes to test or monitor AI impacts or risks	PCP & ARO & MMAE
AG5	Establish processes to measure and assess AI risks	PCP & ARO
AG6	Establish processes to mitigate, rectify, or avoid AI risks	PCP & ARO
AG7	Establish processes to achieve an AI objective	PCP & EPAO
AG8	Assess whether an AI objective is achieved	EPAO & MMAE
AG9	Establish processes to test and monitor AI objectives	PCP & EPAO & MMAE
AG10	Establish processes to measure and assess AI objectives	PCP & EPAO & MMAE
AG11	Provide information about a design decision	UOC
AG12	Determine compliance / Align the systems with a specific standard or guideline	PCP & UOC
AG13	Designate a role	ARRA
AG14	Establish a broad (e.g. ethics review board)	ARRA
AG15	Provide employee training / Ensure workers competence	DEC
AG16	Communicate with or inform users or third parties	DC
AG17	Inform staff and employees about the AI policy	PA

Insights The comparison revealed that ALTAI is centred around trustworthy AI issues and principles rather than how to manage trustworthy AI processes and policies within an organisation. In comparison, the draft AI management system standard does not specifically refer to any trustworthy principle, however, it provides a foundation for implementing these principles in an organisation. The two are therefore complementary regarding effective implementation and

assessment of trustworthy AI, with the comparison providing a way to achieve trustworthiness through management system activities.

Table 3 presents the number of ALTAI activities that are mapped into each AI management system activity. It should be noted that the total number indicates the number of times an AI management system activity is individually mapped to ALTAI activities as the mapping between the two is many-to-many. Activities within AI management system that do not have a corresponding ALTAI activity are omitted from the table (8 in total).

As shown in the table, approximately 50 percent (73 of 144) of ALTAI activities refer to risk management which makes the fact that ALTAI adopts a risk-oriented approach towards trustworthy AI clear. The missing management system activities in the table, which are nearly half of total, demonstrates that processes and tasks at a high level of organisational governance and management are not covered in ALTAI.

Table 3. Number of ALTAI activities mapped into each AIMS activity

AIMS activity	AIMS activity (label)	Nos. ALTAI activities
ARO	Address risks and opportunities	73
PCP	Plan and control AI processes	54
EPAO	Establish and plan to achieve AI objectives	44
DC	Determine AIMS communication	22
MMAE	Monitor, measure, analyse and evaluate AI	20
UOC	Understanding organisation and its context	12
DEC	Determine and ensure competence of people affecting AI performance	7
ARRA	Assign roles, responsibilities and authorities	2
PA	Promote awareness	2

4 Comparison of AI Act with ISO/IEC 42001

4.1 The AI Act Activities

In April 2021, the European Commission published the proposal for EU AI regulation, called AI Act, to create a legal framework for trustworthy AI by laying down obligations which are proportionate to the level of risk imposed by AI systems. Under the AI Act, providers of high-risk AI systems, i.e. systems that are likely to cause harm to health, safety, and rights of individuals, are required to implement a quality management system (Art. 17), among other requirements. The AI Act relies on creation of harmonised AI standards to facilitate conformity to its requirements by providing technical solutions (Art. 40).

Conformity with the AI Act's high-risk AI obligations requires performing organisational as well as technical activities. By analysis of the requirements for

high-risk AI systems and the obligations of providers of those systems, described in title III, Chapters 2 and 3, we identified 52 high-level organisational activities that are `associatedWith` high-risk AI providers, which are modelled as `Agent`s. It is important to note that our list of activities is not exhaustive, and therefore performing the identified activities is essential for conformity to the AI Act but not necessarily sufficient.

4.2 AI Act - ISO/IEC 42001 Activity Comparison

Using the methodology described earlier, we mapped the activities identified from the AI Act to the ones extracted from ISO/IEC 42001. Table 4 shows mapping of AI Act's risk management activities into AI management system.

Table 4. Comparison of AI Act's risk management activities with AIMS

AI Act risk management activity	partOf (AIMS)
Establish risk management system	DC & EIMI & ARO
Implement risk management system	EIMI & ARO
Document risk management system	EIMI & ARO & CUCD
Maintain risk management system	EIMI & ARO
Identify/ Analyse/ Evaluate/ Mitigate Risks	ARO
Communicate Residual Risk to Users	PA & AIRO
Identify Impact On Stakeholders (e.g. children)	USNE & ARO

Insights Our analysis indicates activities to establish management systems, address risks, create documentation, and communicate with external entities are among the most mapped management system activities. This shows that in conformity to the AI Act's legal requirements, documentation and sharing information with external stakeholders are as important as conducting risk management.

Identification of the degree to which compliance to ISO/IEC 42001 assists in conformity to AI Act's high-risk AI obligations needs further investigation as our focus was primarily on the organisational activities explicitly referenced therein.

5 Semantic Modelling of Activities

Documents that specify guidelines generally refer to activities and processes across three distinct phases: ex-ante where a plan of activity must exist; ongoing or during where an activity is currently in the process of being executed; and ex-post where an activity has finished execution or has produced artefacts. For AI guidelines, it is important to model the corresponding semantic representation of activities in a similar manner so as to distinguish when an organisation or system must have a plan in place representing some *future activity* versus having carried

out that activity i.e. *in the past*. This notion is also applicable and demonstrated in the area of legal and regulatory compliance where an obligation can entail provenance of both a plan as well as executed activities, and therefore requires documentation at both ex-ante and ex-post phases [5].

Intended for self-assessment purposes, ALTAI predominately refers to the ex-post phase. This means that to provide answers to ALTAI questions we have to look into the results and artefacts of executed activities. Furthermore, separation between ex-ante and ex-post phases of ALTAI activities enables ex-ante planning for trustworthiness and ex-post trustworthy AI (self-) assessment as outlined by AI management system activities. However, for semantic representation of the activities extracted from ALTAI both planning and execution phases should be taken into account. For example, from 'establish processes to assess AI risks' two activities are inferred: plan for AI risk assessment (ex-ante) and AI risk assessment (ex-post). A semantic model of the former should be able to represent plans for risk assessment, intended steps and actions, responsible parties, and entities generated and used during the planning. This can be done by extending the Ontology for Provenance and Plans (P-Plan)[4]. Naja et al. [6] have adopted the same approach for recording accountability plans. Representing ex-post activities is possible by extending the PROV-O ontology.

To model previously introduced alignment groups we consider the ex-post phase. Each alignment group can be represented as an ontology design pattern (ODP) [7]. An example of one such pattern for AG17 (providing training for employees to ensure competence) that uses the PROV-O ontology to represent agents[5] and activities is shown in fig. 3. By modelling training activities using this pattern all processes and activities which are part of DEC (Determine and ensure competence of people affecting AI performance) can be uniformly represented, and retrieved e.g. using SPARQL queries.

Using the pattern as a generic template for different activities and roles regarding training enables a uniform mechanism to answer questions such as:

- Did the organisation provide training to staff on risk management?
- Who provided the training? When? To whom? On what topic?
- What activities are relevant to training?
- What are the subjects that the organisation provides training on?
- Who is trained on a specific topic, e.g. risk management?

6 Related Work

Boer et al. [8] used an ontology-based approach to facilitate comparison of similar regulations, i.e. in a specific area such as tax, within different jurisdictions. Despres and Szulman [9] proposed an approach for integrating ontologies created from the European community directives. Fiorentini et al. [10] proposed an

[4] `https://www.opmw.org/model/p-plan/`

[5] The PROV concepts of agents and entities are different from ALTAI and AIMS. In PROV, an entity is an artefact such as an input to an activity, and an agent is what is referred to as an entity within ALTAI, AIMS, and the general use of the words.

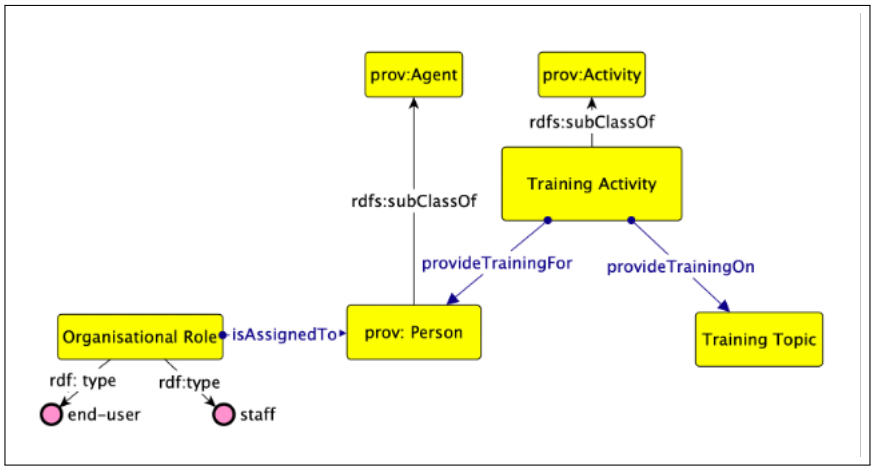

Fig. 3. Training activity pattern.

approach for harmonisation which compares documents using informal analysis, typology of standards, use-cases, and ontologies. Pardo et al. [11] created H2mO - an ontology for harmonisation of reference models and standards utilised in software process improvement. Koelle et al. [12] proposed a tool for ATM security which harmonises relevant standards and regulations. Lewis et al. [3] presented an analysis of the normative content of trustworthy AI guidelines presented by IEEE, EU HLEG, and OECD and mapped these guidelines into ISO 26000 social responsibility issues.

7 Conclusion

This paper presented a comparison and analysis between the EU AI Act, ALTAI, and ISO/IEC AI management system standard to identify the potential alignment between these 3 key documents. The assessment compared management-level activities mentioned in the documents and is represented formally using the trustworthy AI upper-level ontology proposed by [3].

Implications of Comparison and Analysis of AI documents Identification of the gaps existed in the AI documents being developed assists standardisation bodies in determining the areas that need creation or modification of standards. Legislators can use the comparison to determine the degree to which compliance with existing AI standards contributes to conformity to legal obligations and identify the aspects of trustworthy AI that are not subject to regulation. Furthermore, comparison of activities provides a baseline for the communications between authorities and standardisation bodies for development of harmonised regulations and standards.

The comparison assists AI providers and developers in adoption of standards and guidelines required for satisfying legal requirements by helping them identify inconsistencies and areas of overlaps. It can also be used to ensure organisational AI policies are effective in satisfying normative and legal requirements.

Given the potential of AI research to cause harm, recently some AI conferences, such as NeurIPS[6], provide ethical guidelines and ask researchers to assess the impact of their work on key areas of concern, e.g. safety, fairness, and privacy. The comparison methodology can be applied in assessing the alignment of ethical guidelines provided by different conferences, universities' policies on ethics and data protection as well as ethical assessment approaches.

Further Work The comparison presented in this paper will be expanded to provide a more comprehensive analysis and alignment of key terms, technical activities, and requirements detailed within AI documents. Starting with the analysis provided in this paper, we aim to identify a common set of AI risk and impact assessment activities from the AI Act, ALTAI, and ISO risk management and management system standards and extend AIRO - an ontology for describing AI risks [13], to represent provenance of activities. Future work also includes updating this work based on changes made in the subsequent drafts and finalisations of the AI Act and ISO/IEC 42001 standard.

References

[1] *Artificial Intelligence Act: Proposal for a regulation of the European Parliament and the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. 2021. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELLAR:e0649735-a372-11eb-9585-01aa75ed71a1.

[2] European Commission, Content Directorate-General for Communications Networks, and Technology. *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment*. Publications Office, 2020. DOI: doi/10.2759/002360. URL: https://data.europa.eu/doi/10.2759/002360.

[3] Dave Lewis, David Filip, and Harshvardhan J. Pandit. "An Ontology for Standardising Trustworthy AI". In: *Factoring Ethics in Technology, Policy Making, Regulation and AI*. Ed. by Ali G. Hessami and Patricia Shaw. Rijeka: IntechOpen, 2021. Chap. 5. DOI: 10.5772/intechopen.97478. URL: https://doi.org/10.5772/intechopen.97478.

[4] European Commission and Directorate-General for Communications Networks, Content and Technology. *Ethics guidelines for trustworthy AI*. Publications Office, 2019. DOI: doi/10.2759/346720. URL: https://data.europa.eu/doi/10.2759/346720.

[6] NeurIPS 2022 ethics guidelines https://neurips.cc/public/EthicsGuidelines

[5] Harshvardhan J Pandit, Declan O'Sullivan, and Dave Lewis. "Test-driven approach towards GDPR compliance". In: *International Conference on Semantic Systems*. Springer. 2019, pp. 19–33.

[6] Iman Naja et al. "A semantic framework to support AI system accountability and audit". In: *European Semantic Web Conference*. Springer. 2021, pp. 160–176.

[7] Aldo Gangemi. "Ontology design patterns for semantic web content". In: *International semantic web conference*. Springer. 2005, pp. 262–276.

[8] Alexander Boer, Tom van Engers, and Radboud Winkels. "Using ontologies for comparing and harmonizing legislation". In: *Proceedings of the 9th international conference on Artificial intelligence and law*. 2003, pp. 60–69.

[9] Sylvie Despres and Sylvie Szulman. "Merging of legal micro-ontologies from european directives". In: *Artificial Intelligence and Law* 15.2 (2007), pp. 187–200.

[10] Xenia Fiorentini et al. "Towards a method for harmonizing information standards". In: *2009 IEEE International Conference on Automation Science and Engineering*. IEEE. 2009, pp. 466–471.

[11] César Pardo et al. "An ontology for the harmonization of multiple standards and models". In: *Computer Standards & Interfaces* 34.1 (2012), pp. 48–59.

[12] Rainer Koelle, Walter Strijland, and Stefan Roels. "Towards harmonising the legislative, regulatory, and standards-based framework for ATM security: developing a software support tool". In: *2013 International Conference on Availability, Reliability and Security*. IEEE. 2013, pp. 787–793.

[13] Delaram Golpayegani, Harshvardhan J. Pandit, and Dave Lewis. "AIRO: An Ontology for Representing AI Risks Based on the Proposed EU AI Act and ISO Risk Management Standards". In: *Towards a Knowledge-Aware AI* (2022). Publisher: IOS Press, pp. 51–65.