



## Approaching the AI Act... with AI: LLMs and knowledge graphs to extract and analyse obligations<sup>☆</sup>

Federico Galli<sup>a</sup>,<sup>\*</sup> Thiago Raulino Dal Pont<sup>a</sup>, Galileo Sartor<sup>b</sup>, Giuseppe Contissa<sup>a</sup>

<sup>a</sup> University of Bologna, Italy

<sup>b</sup> Swansea University, United Kingdom

### ARTICLE INFO

Dataset link: [https://github.com/thiagordp/obligation\\_extraction\\_for\\_compliance](https://github.com/thiagordp/obligation_extraction_for_compliance)

#### Keywords:

Legal compliance  
Obligation extraction  
Natural Language Processing  
Large Language Models  
Knowledge graphs  
AI act

### ABSTRACT

The EU Artificial Intelligence Act (AIA) exemplifies the growing complexity of digital regulation in the domain of computer technologies. Characterised by abstract terminology, multi-layered provisions, and intersecting regulatory requirements, the AIA poses significant challenges for the identification and interpretation of legal obligations, making compliance a demanding and potentially error-prone endeavour for legal professionals and organisations alike.

Recent advances in Artificial Intelligence (AI), particularly in the fields of Natural Language Processing (NLP) and Large Language Models (LLMs), offer promising support for addressing these challenges. By automating the extraction and structuring of legal rules, AI-based tools have the potential to assist regulatory compliance activities and provide more systematic insights into complex legislative frameworks.

This paper presents an experiment combining NLP techniques and LLMs to automate the extraction and structuring of legal obligations from the AIA.

The approach is based on a modular workflow comprising four main stages: identification of obligations, filtering of deontic statements, analysis of deontic content, and the construction of searchable knowledge graphs. The experiment employed the LLaMA 3.3 70B model, supported by more traditional NLP tools.

Five experts (4 Ph.D. students and 1 post-doc in legal informatics and philosophy) evaluated the system's performance on a subset of cases. The results indicate a precision of 93% in the obligation filtering phase and over 99% accuracy in classifying obligation types, addressees, and predicates. A quantitative analysis of the extracted and analysed obligations revealed a predominance of prescriptive obligations (603 out of 729 total), among which 136 are imposed on the European Commission, while 88 consist of informative duties. The results are in line with current discussions around the AI Act regulatory approach.

These findings underscore the potential of LLM-based tools to enhance regulatory compliance and analysis. Future research will focus on extending the system to additional EU regulations and integrating formal ontologies to enable more advanced representations of legal obligations.

### 1. Introduction

The Artificial Intelligence Act (AIA)<sup>1</sup> currently represents the most ambitious efforts to regulate artificial intelligence (AI) technologies in the global landscape.

The regulation aims to create a harmonised legal framework that ensures AI systems used in the EU are safe, transparent, and accountable. It follows a risk-based approach [1], categorising AI applications based on their potential impact on fundamental rights and safety, with obligations scaling according to the level of risk. For instance, AI systems classified as having “high-risk” must meet strict requirements

<sup>☆</sup> This article is part of a Special issue entitled: ‘Legal Systems’ published in Computer Law & Security Review: The International Journal of Technology Law and Practice.

<sup>\*</sup> Corresponding author.

E-mail addresses: [federico.galli7@unibo.it](mailto:federico.galli7@unibo.it) (F. Galli), [thiago.raulino@unibo.it](mailto:thiago.raulino@unibo.it) (T.R.D. Pont), [galileo.sartor@swansea.ac.uk](mailto:galileo.sartor@swansea.ac.uk) (G. Sartor), [giuseppe.contissa@unibo.it](mailto:giuseppe.contissa@unibo.it) (G. Contissa).

<sup>1</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

<https://doi.org/10.1016/j.clsr.2025.106230>

for data quality, transparency, and human oversight, while “lower-risk systems” are subject only to disclosure obligations.

While the AIA aims to promote responsible AI governance, it also imposes a significant compliance burden on organisations, individuals, as well as EU and national institutions. The duties set out in the Regulation are diverse, covering areas such as data management, risk assessment, governance measure, and procedural safeguards, and apply to a wide range of actors, including AI providers, importers, deployers, Member States, national authorities and EU institutions. For each of these actors, the requirements are numerous and complex, often requiring additional interpretation and implementation steps to achieve compliance. Some obligations may even overlap, such as the transparency obligation for high-risk AI providers (Article 26(7)) and the disclosure requirement for low-risk AI systems (Article 50). In other cases, obligations may entail potentially conflicting compliance measures, such as where providers must ensure both the accuracy of high-risk AI systems (Article 15(1)) and their transparency (Article 13(3)).

To address the challenge of regulatory compliance, there is a growing interest in automated methods for analysing regulatory texts [2–5]. AI-driven computer systems, increasingly using Large Language Models (LLMs) [6], can assist in the compliance-oriented interpretation of regulatory texts. For example, these systems can identify rights, obligations and other critical compliance elements directly from the text [7,8]. They can also link different documents, helping organisations to identify relationships between different regulations, standards and internal policies, and monitor regulatory updates, enabling timely adaptation to new or amended regulations [9]. In addition, AI systems can help interpret ambiguous or complex regulations, provide insight into potential compliance strategies and identify risk areas [10]. The streamlining of the compliance process is said to reduce the time and resources required to interpret and implement regulatory requirements.

Among other things, the ability to automatically extract, detect and classify obligations is crucial for organisations, as it enables them to determine the specific actions or changes they need to implement to comply with the law.

Accordingly, this paper’s research question can be stated as follows: *“To what extent can AI-driven systems, particularly those employing large language models, facilitate the automated extraction and classification of obligations within the context of the EU AI Act (AIA)?”*

The present paper presents an experiment combining NLP techniques and LLMs to automate the extraction and structuring of legal obligations from the AIA.

The paper is structured as follows. Section 2 introduces the relevant previous research and presents how ours builds from those. Then, Section 3 presents the legal theoretical framework backing the research. The processing workflow is presented in Section 4, while Section 5 explains the validation protocol applied to ensure the validity of our results. Finally, Section 6 discusses the results of the experiment and provides a quantitative analysis. Section 7 ends the paper by presenting the conclusions and future research perspectives.

## 2. Related work

Several studies have employed traditional NLP techniques to detect and classify deontic modalities in regulatory texts. Kiyavitskaya et al. [2] developed Gaius T., a tool extending the Cerno framework [11], which used context-free grammars and semantic annotations to extract rights, obligations, actors, and constraints from texts such as the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule and the Italian Stanca Act. Despite the efficacy of Gaius T., challenges were encountered in accurately correlating constraints to subjects within intricate regulatory frameworks.

Similarly, Wyner and Peters [3] proposed a rule-based method to extract conditional and deontic rules, utilising the GATE framework and Stanford Parser to identify legal components such as antecedents

and agents, which are pivotal for interpreting obligations. Dragoni et al. [12] introduced a framework that integrated syntax- and logic-based rule extraction. The approach utilised the Stanford Parser, the Boxer framework, and a lightweight ontology to identify obligations, permissions, and prohibitions. The authors evaluated the approach over the provisions stipulated within the Australian Telecommunications Consumer Protections Code. The study yielded favourable outcomes, with the method demonstrating an accuracy of 80.49% and a recall of 91.67% when extracting the rules contained within the code.

Machine learning (ML) has facilitated the development of more flexible approaches to the obligation extraction task. For example, Waltl et al. [13] addressed the classification of deontic statements in German tenancy law, identifying 22 categories of statements, including prohibitions and permissions, through active learning with multinomial Naïve Bayes, Logistic Regression and Multi-Layer Perceptron classifiers on a corpus of 504 sentences. Gao and Singh [4] also employed ML to classify six types of normative relationships, including prohibitions, authorisations, sanctions, commitments, and powers. The study highlighted the potential of combining syntactic and logic-based NLP methods for effective rule extraction from legal texts.

Recent studies have demonstrated a shift towards neural networks and LLMs. For instance, Neill et al. [5] employs a data-driven methodology to categorise deontic modalities (notably obligations, prohibitions and permissions) in the context of legal language, with a particular emphasis on financial regulations. In this scenario, a more recent study is provided by Liga and Robaldo [14] which, based on previous research [15], explored the use of GPT-3 [16] for Legal Rule Classification (LRC) by fine-tuning the model on a specialised dataset of 707 legal provisions from the GDPR. By combining symbolic knowledge from legal XML standards (LegalDocML and LegalRuleML) with GPT-3’s capabilities, the study aimed to classify legal rules such as obligations, permissions, and constitutive rules. The results showed that GPT-3 significantly outperformed previous models like BERT, achieving a weighted F1 score of 0.93 in the most complex classification scenario, demonstrating its effectiveness even with limited data.

Our paper builds on these advancements by addressing key gaps in the existing literature and experiment a novel approach that leverages LLMs to extract and analyse legal obligations from regulatory text and represent them through knowledge graphs.

Unlike traditional rule-based systems, which rely on predefined rules and struggle to generalise across diverse legal texts and constructs, our work employs LLMs, which are capable of capturing nuanced context and generalising effectively across varied regulatory frameworks.

Similarly, while earlier studies demonstrated high precision in obligation extraction, they often relied on extensive manual annotation to train models, which limited scalability. In contrast, we address this limitation by leveraging pre-trained LLMs, reducing the dependency on expensive and time-consuming annotation processes while maintaining high accuracy in identifying and structuring obligations.

Furthermore, our experiment extends beyond obligation identification by introducing structured analysis rooted in legal theory, which conceptualises obligations as deontic constructs with recurring elements (e.g. addressees, beneficiaries, conditions, etc.). In compliance, these elements are essential to understand the scope of applicability of an obligation and to determine what actions must actually be taken to fulfil it.

Finally, our approach includes a representation of these obligations through knowledge graphs. These graphs enable the visualisation of relationships between obligations, actors, and conditions, supporting a more intuitive navigation of regulatory texts. This structured approach also facilitates advanced functionalities, such as cross-referencing provisions across regulations, network analysis of norms to understand their interdependencies, and logic-based reasoning to identify overlaps, redundancies, or potential conflicts.

### 3. Theoretical framework

Our experiment draws on the work of legal theorists and analytical philosophers who have contributed to the analysis of normative language and the logical structure of legal obligations [17–19]. In particular, the theoretical framework used in our experiment considers the following elements: (1) deontic modality, (2) addressee, (3) predicate, (4) target, (5) specifications, (6) pre-conditions, (7) beneficiary.

We acknowledge that other important elements may define the normative content of obligations. These include, among others, temporal constraints, which establish the time frames within which obligations must be fulfilled, and logical connectors, which clarify the relationships between different components of an obligation (e.g., “AND”, “OR”, “IF-THEN”). Modelling and extracting this information is particularly challenging in the context of our framework [20,21]. As detailed below, the experimented workflow is designed to be modular. Therefore, in the future, it can be easily extended to include other or different elements of deontic obligations.

Elements can be classified as mandatory (non-optional) or optional. The sole compulsory element is the predicate that stipulates the obligation’s content, that is, a state of affairs to be realised or an action to be performed. The remaining elements, including the addressee, the object, the beneficiary, and the pre-condition, are optional. It should be noted that the *addressee* is not mandatory in our representation, as it is not always explicit in legal texts. However, it can be argued that one may always identify it from contextual reading or legal interpretation.

In general, the elements of an obligation may be explicitly or implicitly defined in the text, requiring interpretation or contextual overview to be identified. For instance, even in the absence of explicit mention of a beneficiary in an obligation, this element can usually be inferred from related provisions or the broader intent of the Regulation.

#### 3.1. Deontic modality

Obligations in English-written legal texts are typically conveyed through modal verbs such as “shall”, “must”, “ought to” or “has/have to”. For instance, EU regulation usually employs “shall” to express mandatory actions or states, typically prescribing what must or must not be done in legal and regulatory contexts [22,23].

At the same time, markers, such as “shall”, may exhibit a high level of semantic ambiguity. For example, research in legal-linguistic has noted that regulatory statements introduced by “shall” may refer to different meanings [24,25]. For example, “shall” may be used to provide a definition or explanation of a certain concept or entity, as in “any reference to an economic operator under Regulation (EU) 2019/1020 shall be understood as including all operators identified in Article 2(1) of this Regulation” (Article 74 AIA). Other times, “shall” may be used as a constitutive statement, i.e. those creating institutional entities by assigning them specific statuses and/or functions [26], as in “The Board shall be composed of one representative per Member State.” (Article 65(2) AIA). Disambiguating such instances is particularly relevant to extracting deontic obligations correctly. For this reason, as we will detail in Section 4, we have implemented an obligation filtering step to isolate genuine deontic obligations.

The concept of obligation also includes prohibitions, as a prohibition of a certain action is assumed to be equivalent to the obligation to refrain from performing that action. For instance, Article 5(1)(a), which prohibits AI systems from being used for the purpose or with the effect of manipulating individuals and thereby causing them harm, can be interpreted as implying an obligation to refrain from such practices. This correspondence is also reflected at the linguistic level, as most prohibitions are denoted by the negative form of the verbs establishing obligations (e.g., “must not” or “shall not”). Other textual indicators of prohibition may include phrases like “is forbidden”, “is prohibited”, or “is not allowed”, which help clarify the restricted nature of certain actions.

Conversely, the deontic modality addressed in this paper does not include rights. However, as noted by Hohfeld [17], granting rights means imposing obligations towards the rights-holder. An aspect that will be considered in the future is exactly the interplay between rights and obligations.

As a last note, our framework distinguishes between Obligations of Action and Obligations of Being [27]. Obligations of action require an action to be taken in order to achieve compliance. The level of specificity for the required action can vary. For example, compare these two provisions: “The AI provider shall implement appropriate safety measures” and “The data controller shall maintain a record of processing activities”. Despite the difference in detail, the key requirement remains the same: an action, whether clearly defined or broadly described, is explicitly mandated.

In contrast, Obligations of Being demand that an entity (a thing, a person, an organisation) meets certain requirements. For instance, an AI system may be required to be secure by design, which implies meeting security standards but does not specify exact security measures.

This distinction applies equally to prohibitions, as prohibitions can either prevent specific actions or impose constraints on the state of being. For instance, consider a regulation stating that “providers shall not use AI systems to exploit vulnerabilities of specific groups” and that “AI systems shall not contain biases in their design resulting in unfair treatment of protected groups”. In the first case, the regulation would introduce an actionable restriction, requiring the avoidance of a specific design practice. In the second case, it would determine a desired state of conformity, imposing a constraint on the attributes of the regulated target (the AI system).

It can be argued that any obligation/prohibition of being can ultimately be translated into an obligation/prohibition of Action, once the addressees, i.e. the entities responsible for implementing the requirements, are identified. Nevertheless, it is important to note that legal texts do not always explicitly determine the specific addressee(s) or the specific actions to be undertaken for that purpose. This determination can be drawn contextually upon reading other provisions in the same text, or implicitly by using interpretative means.

#### 3.2. Predicate

The predicate defines the required state or action associated with the obligation. In Obligations of Being, the predicate specifies the requirement that the object should satisfy and is often expressed in the passive voice (e.g. “shall be resilient”). In obligations of action, the predicate describes the specific action required and can be in either the active or passive voice (e.g. “shall perform”, or “must perform”). Irrespective of its form, the predicate conveys the core content of the obligation, clarifying either the state in which the object is to be or the action that must be performed concerning the object.

#### 3.3. Addressee

This element represents a core aspect of an obligation. An addressee is defined as the individual, organisation, or entity bearing active responsibility for either realising, maintaining or affecting the target’s condition (in Obligations of Being) or performing the specified action (in Obligations of Action). The addressee must individually or collectively possess the capacity to act, such as a natural or legal person. Consequently, entities like “AI systems”, “a process”, or “adopted measures” cannot serve as addressees of obligations.

In Obligations of Action, the addressee is generally explicitly stated in the regulatory provisions. At the same time, there may be cases where only the target and the predicate (generally expressed in the passive form) are explicit, and the addressee who has to carry out the action is implicit, requiring interpretative analysis to identify who must ensure the specified condition is met. For example, in obligations of

action, the responsibility to conduct a particular task (like a risk assessment) is typically assigned directly to an entity such as a “provider” or “organisation”.

In contrast, Obligations of Being require a regulated target to achieve certain standards without directly naming one or more addressees. Consider the provision that “the oversight measures shall be commensurate with the risks” (Article 14 of the AIA). In this case, the Regulation leaves it to the reader to infer who is responsible for ensuring that the oversight measures meet the required standard. This is often based on the broader regulatory context or the nature of the entity most logically associated with the target (i.e., the deployer of the AI system).

### 3.4. Target

The target is the entity associated with the predicate, defining what must be realised, maintained or affected. It is an optional element. The role of the target differs depending on the obligation type: for Obligations of Being, the target is something that must meet or uphold a specific state, quality, or condition (e.g., “a system must remain secure”). Here, the target is always the grammatical subject of the sentence. For Obligations of Action, the target is the entity that is subject to or concerned by the action and is the grammatical object of the sentence. When the predicate consists of a ditransitive verb (e.g., inform, grant, notify), the target is always the direct object of the sentence, while the indirect object is the beneficiary (see below). For instance, in the statement “information on fines imposed under this Article shall also be communicated to the Board as appropriate” Article 101(4), the target is “information on fines imposed under this Article” while the Board is the beneficiary.

### 3.5. Specification

A specification defines the standard, time or methods required to fulfil the obligation. For Obligations of Being, they typically establish qualitative or quantitative criteria specifying the requirements the target is expected to meet (e.g., “in accordance with the standard procedure”). In obligations of action, specifications may outline the rigour or procedural benchmark necessary to perform the action. Not all obligations are accompanied by specifications; therefore, this element is optional.

### 3.6. Beneficiary

The beneficiary is the party who directly or indirectly benefits from the fulfilment of the obligation.<sup>2</sup> In Obligations of Being, the beneficiary may benefit from the maintained condition or state of the target, such as enhanced safety or transparency. In obligations of action, the beneficiary derives an advantage from the specific action performed, such as receiving a service, right, or information. While the beneficiary is not always explicitly stated, it is often possible to infer their identity from the context or prior knowledge about the relationships and responsibilities between the involved parties.

In theory, when a beneficiary can be identified, it should always be possible to establish derivative obligations (such as the duty to compensate for harm) that come into effect in the event of a violation of the primary obligation. For now, we do not consider these derivative obligations (also known as “contrary-to-duty obligations” [28]) in our experiments.

<sup>2</sup> An indirect beneficiary refers to a situation where the party benefiting from the obligation differs from the one directly advantaged by its performance. Although this distinction may be relevant in our contexts, so far we have not implemented this distinction in the prompt.

### 3.7. Pre-condition

A pre-condition is an optional element that defines the circumstances or criteria that must be met (or not met) before an obligation becomes enforceable. In both Obligations of Being and Obligations of Action, pre-conditions act as triggering factors. Positive pre-conditions require certain conditions to be fulfilled for the obligation to take effect, while negative pre-conditions specify that the obligation applies only if particular conditions are absent.

In our framework, we also consider exceptions as instances of negative pre-conditions. These exceptions can be identified explicitly in the deontic statement or inferred from its context or referenced provisions. However, we acknowledge that not all exceptions can be modelled and identified as negative pre-conditions of particular obligations. Exceptions may stem from broader contextual factors or require a cross-referenced interpretation of related provisions.

## 4. Processing workflow

Fig. 1 presents the general overview of the experimented workflow.

In particular, the workflow is composed of four modules: Obligations Detection (Module 1); Deontic Obligations Filtering (Module 2), Deontic Obligations Analysis (Module 3); Deontic Obligations Representation (Module 4).

Modules 2 and 3 employ an LLM for the purposes of text classification [29] and information extraction [30] tasks, respectively. In particular, in this study, we experimented with Meta’s LLaMa 3.3 70B, using Together AI API,<sup>3</sup> with default parameters employed and temperature set to 0.7.

We limited the study to a single model for reasons of scope and feasibility. Each sentence required detailed expert annotation across multiple information categories, making large-scale, multi-model testing prohibitively resource-intensive. Moreover, the purpose of this work is to validate the feasibility of a modular workflow for legal text processing, not to perform a benchmarking exercise across models. Within this scope, we selected LLaMA because it is open-source, ensuring transparency and reproducibility of our experiments.<sup>4</sup>

The following subsections provide a detailed description of the workflow.

### 4.1. Obligations detection

The first step of the workflow consists in identifying all sentences within the regulation that may contain deontic obligations.

To this end, we retrieved the full text of the Artificial Intelligence Act from the EUR-Lex database, the official EU repository of legal documents.<sup>5</sup> The database provides legal texts in a partially machine-readable XML format, in which documents are already segmented into paragraphs. This structure facilitates the decomposition of the legal text into manageable units for further computational analysis.

We then proceeded to identify all paragraphs potentially containing deontic obligations. As outlined in Section 3.1, this identification step relies on the presence of linguistic markers commonly used in EU legal drafting in English, such as “shall”, “must”, “should”, “has”, and “have to”.<sup>6</sup> For each relevant paragraph, we extracted the specific sentence in which the keyword appears. This ensures the precise isolation of

<sup>3</sup> <https://api.together.ai/>

<sup>4</sup> While larger proprietary models (e.g., GPT-5, Gemini, Claude) may achieve higher accuracy, our choice of LLaMA provides a conservative baseline. If an open model can achieve the reported results, stronger proprietary models are expected to perform at least as well. This strengthens the generalisability of the findings without relying on closed systems.

<sup>5</sup> <https://eur-lex.europa.eu>

<sup>6</sup> Following Sandrelli et al. [22] and the English translation guide from European Commission [23].

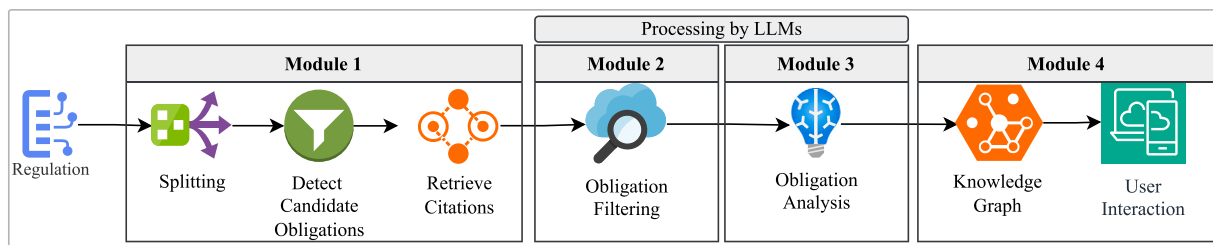


Fig. 1. Workflow for obligations processing comprising four modules: detection, filtering, analysis and knowledge graph representation. The workflow leverages LLMs to filter and analyse provisions in order to detect and structure deontic obligations for user interaction.

the statement potentially containing deontic obligation, which is then submitted for further analysis by the LLM in the following steps.

While modal verbs such as *shall* and *must* capture the majority of obligation statements, we acknowledge that a purely keyword-based filtering approach may miss obligations expressed through alternative formulations. To partially address this risk, our workflow included manual random checks and cross-linguistic comparisons.<sup>7</sup>

To preserve the necessary legal context for accurate analysis, we also retained supplementary information.

First, we included the full text of the paragraph where the statement potentially containing a deontic obligation appears. This broader context is often essential, as it may contain key elements — such as addressees, specifications, or preconditions — referred to indirectly through pronouns (e.g., “This shall...”) or intra-textual references (e.g., “as stated in the previous paragraph”).

Second, we incorporated the text of any additional AI Act paragraphs explicitly cited within statements potentially containing deontic obligations. We hypothesised that these cross-referenced provisions may also provide relevant information that enhances the LLM’s ability to perform semantically correct analysis.

#### 4.2. Deontic obligations filtering

In this module, the LLM analyses the sentence potentially containing deontic obligations to retain only those that actually contain one or more deontic obligations. As previously expressed, not all sentences containing modal verbs (e.g., “shall”, “must”) convey a deontic meaning. Some provisions may express *definitions*, *authorisations*, or *constitutive statements*. Consequently, this filtering step identifies deontic obligations and prohibitions. To do so, we prompted the LLM to classify a sentence potentially containing deontic obligations in one of the below categories (see [Appendix A](#) for details on the prompt engineering).

- **Deontic Obligation:** The statement imposes a duty on someone (Obligation of Action) or establishes a requirement on something (Obligation of Being), e.g., “Importers shall indicate their name, registered trade name or registered trade mark, and the address at which they can be contacted on the high-risk AI system and on its packaging or its accompanying documentation, where applicable.”(Article 23(3)).
- **Definition:** The statement provides a legal definition of an entity or term, e.g., “any reference to an economic operator under Regulation (EU) 2019/1020 shall be understood as including all operators identified in Article 2(1) of this Regulation.” (Article 74(1)).

- **Constitutive Statement:** The statement creates a new state of affairs or qualifies a fact with a legal effect, rather than prescribing behaviour or imposing obligations, e.g., “in addition to the high-risk AI systems referred to in paragraph 1, AI systems referred to in Annex III shall be considered to be high-risk.” (Article 6(2)).
- **Entitlements:** The statement empowers someone or grants a right, e.g., “only the Commission and national authorities referred to in Article 74(8) shall have access to the respective restricted sections of the EU database listed in the first subparagraph of this paragraph.” (Article 49(4), last period).
- **Authorisations:** The statement indicates when an action is permitted or authorised under specific conditions, e.g., “This shall not preclude the use of assessed high-risk AI systems that are necessary for the operations of the conformity assessment body, or the use of such high-risk AI systems for personal purposes.” (Article 31(4), last period).
- **Deontic Prohibition:** The statement prohibits someone (Obligation of Action) or establishes a negative requirement on something (Obligation of Being), e.g., “They shall neither seek nor take instructions from anyone when exercising their tasks under paragraph 3.” (Article 68(4), second period).

An additional label, “Not Applicable”, was introduced to serve as a default category for statements that do not fall within any predefined obligation types. This label also functions as a safeguard against false positives that may have passed through the keyword-based obligation detection phase into the deontic obligation filtering step.

For the output, the LLM was tasked to provide a classification and a justification to each candidate sentence, based on its content and the content of its paragraph. An example of this output, for a sentence from AIA Article 61(2), is demonstrated in Listing 1.

**Listing 1** Example of LLM output for obligation filtering from AIA Article 61(2) (“The informed consent shall be dated and documented and a copy shall be given to the subjects of testing or their legal representative.”).

```

1 {
2   "classification": "Deontic obligation",
3   "justification": "This statement imposes
a duty on the person or entity
responsible for obtaining informed
consent to perform specific actions:
dating and documenting the consent,
and providing a copy to the subjects
of testing or their legal
representative. It regulates their
behaviour by mandating these
activities."
4 }

```

<sup>7</sup> For example, we compared selected English provisions with their counterparts in the French and German versions of the AI Act to verify whether obligations expressed without modal verbs such as “shall” or “must” were being overlooked. Such an analysis did not reveal any substantial omissions.

### 4.3. Deontic obligations analysis

Based on the results of the obligation filtering step, the LLM<sup>8</sup> is required to process each deontic obligation and analyse them in further detail. Following our theoretical framework (Section 3), the task is first to classify the type of deontic obligation (i.e. obligation of action/obligation of being) and, for each of them, correctly identify the six key elements where present. These elements comprise the *addressee*, *predicate*, *targets*, *specifications*, *preconditions*, and *beneficiary*.

The LLM was assigned the task of analysing the deontic obligation and extracting the aforementioned elements (see Appendix B for details on the prompt engineering).

As an additional instruction, we provided guidelines for determining the extraction method, that is, identifying the source from which each element is derived. The following extraction methods are presented:

- *Stated*, when the information is explicitly present in the analysed deontic obligation;
- *Context*, when it is inferred from the surrounding text;
- *Citation*, when it is inferred from referenced sections;
- *Background Knowledge*, when it is inferred based on the model's prior knowledge;
- *None*, when no reliable extraction is possible.

We included this additional classification task for the LLM for both technical and legal reasons. Technically, initial results showed that prompting the LLM to classify the extraction methods minimised the risk of hallucinations, especially when the information was not explicitly stated in the provision. From a legal standpoint, this approach ensures a comprehensive understanding of obligations by considering not only explicitly stated information but also the broader context and referenced sections. Methods like “Context” and “Citation” enable the reconstruction of obligations dispersed in the surroundings of the analysed provision.

Conversely, the “Background Knowledge” category ensures that when explicit information is unavailable in the provision or its context/extracted references, the model is allowed to make probabilistic inferences based on its pre-existing knowledge. Besides ensuring the completeness of the extraction, this category provided a means to test the model's understanding of general legal knowledge and its capacity for logical reasoning beyond the explicit text.

The model outputs a JSON object containing the relevant elements (addresses, predicate, etc.) in terms of the information itself (value) and its source (extraction method). An example of this output, for a sentence from AIA Article 61(2), is demonstrated in Listing 2.

**Listing 2** Example of LLM output for obligation filtering from AIA Article 61(2) (“The informed consent shall be dated and documented and a copy shall be given to the subjects of testing or their legal representative.”).

```

1  [{"ObligationTypeClassification": "
2  Obligation of Action",
3  "Addressees": [
4    {
5      "extraction_method": "
6      Background-Knowledge",
7      "value": "The provider or its
8      legal representative"
9    }
10 ]
  },
  "Predicate": {
    "extraction_method": "Stated",

```

```

11     "value": "shall be dated and
12     documented",
13     "verb": "passive"
14   },
15   "Targets": [
16     { "extraction_method": "Stated",
17       "value": "The informed consent"
18     }
19   ],
20   "Specifications": [
21     { "extraction_method": "None", "
22       value": null }
23   ],
24   "Pre-Conditions": [
25     {
26       "extraction_method": "Context",
27       "value": "freely-given informed
28       consent shall be obtained
29       from the subjects of testing
30       prior to their participation
31       in such testing"
32     }
33   ],
34   "Beneficiaries": [
35     {
36       "extraction_method": "Stated",
37       "value": "The subjects of testing
38       or their legal representative"
39     }
40   ]
41 },
42 {(Hidden) One more obligation extracted...}
43 ]

```

It may be observed that the content of cited sections and paragraphs has not been included in the Deontic Obligation Filtering task, despite its inclusion in the Deontic Obligation Analysis. The decision was taken to proceed with the aforementioned course of action, as it was observed that there had been no enhancement in the Deontic Obligation Filtering outputs when the citations were given. Conversely, an enhancement in the Deontic Obligation Analysis was noted. Furthermore, the reduction in the size of the LLM input prompt results in a decrease in both computational and financial costs.

### 4.4. Deontic obligations representation

The Deontic Obligation Representation module is the final step of our experimental workflow, where the extracted and structured deontic obligations are represented semantically. In particular, to effectively visualise and enable querying of the extracted obligations, we leverage knowledge graphs, which highlight the relationships between entities, actions, and other elements involved in the obligations.

To construct a unified and semantically rich knowledge graph for legal obligations, an ontology is developed to normalise entities, actions, and relationships. The generalised concepts, previously selected by legal experts, are used to cluster similar terms, ensuring consistency and facilitating semantic searches.

The semantic clustering is supposed to align specific occurrences with the generalised ontology elements, grouping semantically related terms. This approach ensures that variations in syntax or phrasing do not hinder accurate obligation representation. This is carried out using Sentence Transformers, using general models (a smaller one all-MiniLM-L6-v2, and nomic-ai/modernbert-embed-base, a more modern and larger version), as well as models trained on EU legislation<sup>9</sup>

<sup>8</sup> The same setup from the Obligation Filtering step is used.

<sup>9</sup> All the models used were fetched from HuggingFace.

**Table 1**

Evaluation results for the deontic obligation filtering task. *Accuracy* measures the system's performance against the resolved expert judgements. *Krippendorff's Alpha* ( $\alpha$ ) measures the inter-rater reliability among the five human experts to assess the task's subjectivity.

Element	Accuracy	Krippendorff's Alpha ( $\alpha$ )
Classification	0.93	0.29
Justification	0.90	0.26

Once constructed, the final knowledge graph can be exported in GraphML format, allowing it to be visualised or integrated into graph database systems (see Section 6).

## 5. Expert validation protocol

We validated our experiments through a two-stage process in order to ensure robust and impartial evaluation.

Firstly, 60 randomly selected sentences were evaluated blindly by 4 independent legal (Ph.D. students in legal informatics and philosophy). For this purpose, the corpus was divided into two subsets of 30 sentences, and each subset was then reviewed by two experts also independently. The evaluation focused on two key aspects: obligation filtering (correctly identifying whether a sentence contains a deontic obligation) and obligation analysis (the accurate extraction of components such as obligation type, addressee and predicate).

In the second phase of the experiment, a fifth independent legal expert (post-doc in legal informatics and philosophy) was tasked with blindly evaluating the full sample of 60 sentences. This review resolved the disagreements arising from the initial assessments. In this way, each sample of the 60 was rated by three independent experts. Reliability was then quantified as the proportion of sentences for which all three reviewers were in complete agreement (i.e. all marked as correct or all marked as incorrect).

We acknowledge two important limitations of this validation approach. First, the size of the evaluated sample is limited, which constrains the statistical significance and generalisability of the results. Second, the binary "correct/incorrect" design prevents the use of standard classification metrics such as precision and recall, as well as more comprehensive human-human or human-system agreement analyses.

Nevertheless, the evaluation remains meaningful: it highlights the feasibility of the workflow, identifies areas where annotator agreement is more challenging, and provides an initial reproducible baseline for future studies. Extending the validation framework to include larger samples and multi-annotator gold standards allowing richer agreement metrics is identified as an important direction for future work.

## 6. Results and discussion

In this Section we present and discuss the results of our experiments.

### 6.1. LLM evaluation

The outcomes of the Deontic Obligation Filtering task are reported in Table 1 inclusive of the system's performance via accuracy and on the reliability of the human evaluation via Krippendorff's Alpha ( $\alpha$ ) [31].

Table 2 summarises the results for the Deontic Obligation Analysis task, focusing on the levels of accuracy and Krippendorff's Alpha score for the extraction of specific obligation elements (see Section 3). These components include both the extracted values and their identified sources ('extraction methods').

The Deontic Obligation Filtering task demonstrated high performance, achieving 93% accuracy in classifying provisions and 90% accuracy for justifying its classifications. While these accuracy scores are strong, the Krippendorff's Alpha scores indicate only "fair agreement"

**Table 2**

Evaluation results for the deontic obligation analysis task. *Accuracy* measures the system's performance against the resolved expert judgements. *Krippendorff's Alpha* ( $\alpha$ ) measures the inter-rater reliability among the five human experts to assess the inherent subjectivity of each extraction task.

Element	Feature	Accuracy	Krippendorff's Alpha ( $\alpha$ )
Obligation Type	Value	>0.99	1.00
	Extraction Method	>0.99	-0.01
Addressees	Value	0.97	0.47
	Extraction Method	0.97	0.47
Targets	Value	0.89	0.60
	Extraction Method	0.94	0.24
Predicate	Value	>0.99	-0.02
	Extraction Method	>0.99	-0.02
Specifications	Value	0.94	0.03
	Extraction Method	0.94	0.06
Pre-Conditions	Value	0.97	0.06
	Extraction Method	0.97	-0.07
Beneficiary	Value	0.97	-0.03
	Extraction Method	0.97	-0.03

among the five human experts, with an  $\alpha$  of 0.29 for classification and 0.26 for justification. This contrast is insightful in that it suggests that, although the LLM's output aligns well with the *gold standard*, the task of differentiating between deontic obligations and other regulatory statements is inherently subjective and allows for significant variance in interpretation among legal experts.

Similarly, the Deontic Obligation Analysis task demonstrated a high degree of accuracy in most of the elements, especially with regard to categorising obligation types and extracting predicates and addressees.

More specifically, the Obligation Type, Addressees and Predicate values were recognised with over 99% accuracy. The extraction of Specifications and Targets demonstrated comparatively lower accuracy, with Specifications at 94% and Targets identified correctly in 89% of cases. Pre-conditions and Beneficiaries were also identified with very high accuracy (both at 97%).<sup>10</sup>

The Alpha scores, however, reveal a significant variation in the difficulty and subjectivity of the Deontic Obligation Analysis. For the high-level task of identifying the Obligation Type, a perfect score ( $\alpha = 1.0$ ) confirms the task is well-defined and the evaluation is reliable. For more interpretive components like identifying Targets ( $\alpha = 0.60$ ), the score indicates moderate agreement, effectively quantifying the task's subjectivity. Most notably, for fine-grained tasks such as identifying the Predicate ( $\alpha = -0.02$ ) and Beneficiary ( $\alpha = -0.03$ ), the negative Alpha scores signify systematic disagreement. This indicates that the experts were not making random errors but were consistently applying divergent internal criteria to their binary (Correct/Wrong) judgements.

Overall, while the system performs well in isolating the core legal elements (e.g. obligation types, predicates, and addressees), it reveals limitations when dealing with more interpretatively demanding components (e.g. beneficiaries, standards, and preconditions). These findings emphasise the necessity for continued refinement of LLM-based systems, particularly with regard to enhancing their capacity to tackle semantic ambiguities and the more profound inferential structures that are characteristic of legal texts.

### 6.2. Quantitative legal analysis

Based on the results of the filtering and analysis modules on the whole dataset, we performed a quantitative analysis to highlight the

<sup>10</sup> We report accuracy values as ">0.99" instead of "1" to avoid suggesting absolute perfection. Given the limited dataset size, reporting a rounded "1.0" might misleadingly imply flawless performance. The notation ">0.99" indicates extremely high accuracy while maintaining methodological caution.

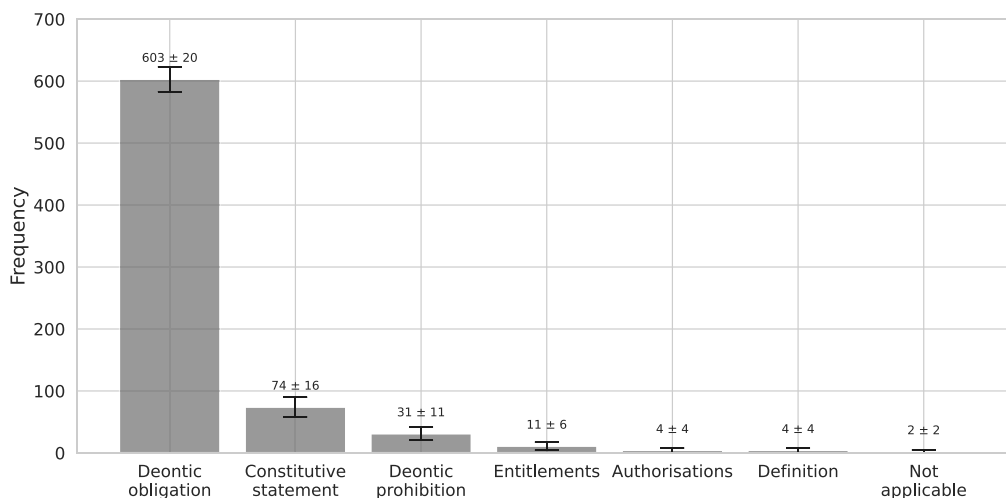


Fig. 2. Obligation filtering: frequency distribution of assigned categories, and statistical analysis with 95% confidence intervals (± CI) shown for each category.

most recurrent types of obligations and addressed entities. Many of the findings align with the current doctrinal analysis on the AI Act and its regulatory approach.

**Deontic obligations.** Fig. 2 shows the distribution of potential obligations along deontic obligations and the other categories pertinent to the Deontic Obligation Filtering module. The results are presented with a 95% confidence intervals.

The results show a strong dominance of deontic obligations (603 ± 20 out of 729 total). This suggests that the provisions in most parts of the AI Act are indeed aimed at compelling various entities to modify their behaviour in accordance with prescriptive regulatory standards. Deontic obligations are accompanied by approximately 31 instances (31 ± 11) of deontic prohibitions, most of which are concentrated in Article 5, which lists the practices explicitly prohibited under the regulation (e.g., manipulative or exploitative AI systems). The disproportionate prevalence of obligations compared to prohibitions reflects a broader regulatory trend within the EU to favour proactive governance through mandated actions over passive restriction.

The second most frequent category, although significantly lower in magnitude, is constitutive statements, which appear approximately 74 times (± 16). These provisions are nonetheless central to the legal architecture of the AI Act, as they serve to define legal statuses and assign roles to various actors (e.g., “the AI Office”, “notified body”, “the Board”), thereby shaping the ontological framework within which deontic norms operate.

Finally, the LLM identified a very limited number of entitlements or authorisations (4 instances, ± 4). This low frequency aligns with the design of the AI Act, which is primarily obligation-centric and does not explicitly aim to confer individual new rights or freedoms (except from those, included in Article 85 and 86 in the very last stages of the proposal) [32]. Rather, its emphasis lies in imposing technical, procedural, and organisational duties on actors involved in the development, deployment, and oversight of AI systems.

At the same time, it is open to interpretation that certain rights may be considered correlative to the obligations imposed on providers and other entities. This means that, while the Act does not directly frame its provisions in terms of subjective rights, individuals or groups may nonetheless derive protections or entitlements from the effective enforcement of those deontic obligations.

**Deontic obligation types.** Table 3 presents an analysis of obligation types derived from the AI Act, categorising them into “Obligations of Action” (754 instances) and “Obligations of Being” (108 instances). The distribution highlights the prevalence of “Obligations of Action”,

**Table 3**  
Distribution of extracted obligation across deontic obligation types.

Obligation Type	Count (± 95% CI)
Obligation of Action	754 ± 54
Obligation of Being	108 ± 20

**Table 4**  
Distribution of extracted obligation across addresses.

Addressee	Count	Addressee	Count
European Commission	146	Authorised Representative	16
Provider	93	Importer	13
Member State	48	Scientific Panel	11
Notified Body	38	Other Body	10
AI Office	36	Distributor	10
Market Surveillance Authority	35	Provider of General-Purpose AI Model	9
National Competent Authority	31	Union Institution	8
Notifying Authority	26	Office	8
Provider of High-Risk AI System	23	Agency	8
Deployer	20	Board	8

indicating that the AI Act focuses significantly on duties that mandate specific actions on relevant entities.

However, the notable presence of “Obligations of Being” points towards an emphasis on the Regulation of maintaining certain requirements or states of being. In the AI Act, this is often, but not exclusively, related to the adherence of risk AI systems with essential requirements or general performance criteria rather than specific, quantifiable actions. This approach aligns with outcome-based regulation, focusing on achieving desired results rather than prescribing fixed methods. It also reflects the EU’s new legislative approach to product safety, wherein broad objectives are established to ensure technology neutrality and adaptability within the AI sector [33].

**Addressees.** Table 4 presents data on the addressees, identifying the entities most frequently designated with compliance duties.

Among specified addressees, institutional actors, such as the European Commission, Member States, AI Office, etc., are prominent. This is in line with the fact that a substantial portion of the Regulation is concerned with establishing a governance framework and oversight mechanisms for AI in the EU [34]. The European Commission, with 146 instances, stands out as a primary actor in the AI Act, reflecting its central role in further implementation, oversight, coordination,

**Table 5**  
Distribution of extracted obligation across predicates.

Predicate	Mentions	Predicate	Mentions
shall inform	44	shall keep	11
shall ensure	37	shall verify	10
shall provide	34	shall facilitate	9
shall take into account	23	shall draw	8
shall include	22	shall evaluate	8
shall adopt	18	shall assess	8
shall cooperate	14	shall be prohibited	8
shall notify	14	shall establish	7
shall not apply	12	shall report	7
shall submit	11	shall contain	7

and enforcement. This is even more so if one considers that the AI Office, which is the Commission’s internal function devoted to such governance activities, is also ranked high with 36 instances.

Among non-institutional entities, the provider is the most frequently cited, with 85 instances, high-lighting the significant compliance responsibilities for those designing and bringing AI products to market. This count increases when considering specific categories, such as providers of high-risk AI systems (23 mentions) and providers of general-purpose AI models (9 instances). Deployers are also frequently mentioned (20 instances), underscoring their role in ensuring AI systems are used in compliance with regulatory standards.

*Predicates.* Table 5 details the number of obligations of action per predicates. This categorisation is relevant as it demonstrates the emphasis on different types of compliance actions the AI Act prioritises for effective governance.

The top-ranked category is “shall inform” with 88 counts, followed by “shall ensure” (74 counts). These two types of predicates are indeed indicative of the regulatory attitude of the AI Act.

First, the duty to inform is indicative of the transparency-based rationale of the AI Act. This is reinforced by other predicates such as “shall notify”, “shall submit”, and possibly also “shall provide”.<sup>11</sup> All these actions highlight an emphasis on the need for clear communication among stakeholders, including both institutional and non-institutional actors, about AI operations [35].

On the other hand, predicates such as “shall ensure” and “shall take into account” (28 counts) reflect a performance-based regulatory approach. By requiring addressees to achieve certain results (“shall ensure”) or considering certain regulatory priorities (“shall take into account”), the AI Act emphasises the responsibility of AI operators to take appropriate steps to comply with the overarching normative requirements without rigidly specifying the exact methods [36]. At the same time, such an approach may also be criticised for introducing ambiguity and a lack of guidance on the specific measures to be adopted, leaving organisations uncertain about the exact requirements needed to meet compliance standards.

*Beneficiaries.* Table 6 lists the most affected beneficiaries, identifying the groups or entities that stand to benefit from compliance with the Regulation. This information is pertinent for clarifying the regulatory intent to protect specific stakeholders and aligns with the broader societal objectives of the AI Act.

As with the addressees, the beneficiary statistics reveal a high frequency of institutional actors as primary beneficiaries. For instance, the European Commission (41 counts), the AI Office (16 counts), and other EU institutions and national authorities appear prominently. In particular, institutional beneficiaries such as the European Commission and Member States highlight the AI Act’s focus on governance, central oversight, and the harmonisation of AI-related standards across the EU.

<sup>11</sup> Based on our experience in reading the AI Act, we may assume that many of these predicates in this class can be related to “information” as an object.

**Table 6**  
Distribution of extracted obligation across beneficiaries.

Beneficiary	Count	Beneficiary	Count
European Commission	41	Notifying Authority	8
Member State	23	Deployers	7
National Competent Authority	22	Small and Medium-Sized Enterprises (SMEs)	6
Provider	17	Start-ups	6
Market Surveillance Authority	16	Prospective Provider	6
AI Office	16	Competent Authority	5
Board	11	Relevant Competent Authority	5
European Parliament	10	Member State Concerned	5
Council	10	Relevant Operator	5
Notified Body	9	Relevant Market	5
		Surveillance Authority	

This focus on institutional beneficiaries implies, again, that the regulation is designed to set a robust governance framework, where oversight bodies are empowered and supported to monitor and control AI development and deployment. In addition, including entities like the National Competent Authority (22 counts) and the Market Surveillance Authority (16) as beneficiaries reflects the AI Act’s commitment to ensuring that regulatory bodies are equipped with the information and resources necessary for enforcing compliance.

Meanwhile, non-institutional beneficiaries, although less frequent, are also acknowledged, including deployers, SMEs, and start-ups. The presence of these beneficiaries indicates that the Regulation also seeks to allocate responsibility along the AI value chain, also to the benefit of smaller actors who contribute to AI development and deployment [37].

### 6.3. Knowledge graphs

All extracted obligations were integrated into a knowledge graph to support the Deontic Obligation Representation task.

The graph contains two main node types, “Entity” and “Person”, with two relationships, “Obligation” and “Has Child”.

The Person node represents Addressees or Beneficiaries as they appear in the source text, while the Entity serves as a semantic cluster encompassing multiple Person elements, thereby consolidating diverse textual references under unified conceptual labels, highlighting both obligations and compositional relationships (as mentioned in 4.4). Directed edges labelled *has\_obligation* denote the obligation as extracted from the pipeline, with properties for the Specifications or Conditions. Finally the Has Child edges indicate the structural or hierarchical grouping of the Person nodes. Using the compositional structure, the graph enables searching not only for Person-Person (Addressee-Beneficiary) obligations, but also inter-entity obligations, abstracting the term from the specific formulation, as can be seen in Fig. 3, where the search is carried out for the generic term “provider”. It is possible to show also the intermediate “Person” nodes to have a more precise view if necessary.

The resulting graph, exported in GraphML format, may facilitate semantic querying, visualisation and compliance checking. Consequently, it enhances the interpretability and navigability of complex regulatory texts.

## 7. Conclusion and future work

The present study presented an experiment to prove the extent to which AI-driven systems, particularly those employing LLMs, can facilitate the automated extraction and semantic classification of legal obligations within the context of the EU AI Act.

The findings of this study demonstrate that LLM-based systems, when supported by legal theory and integrated into a structured workflow, may significantly facilitate the navigation of a complex regulatory

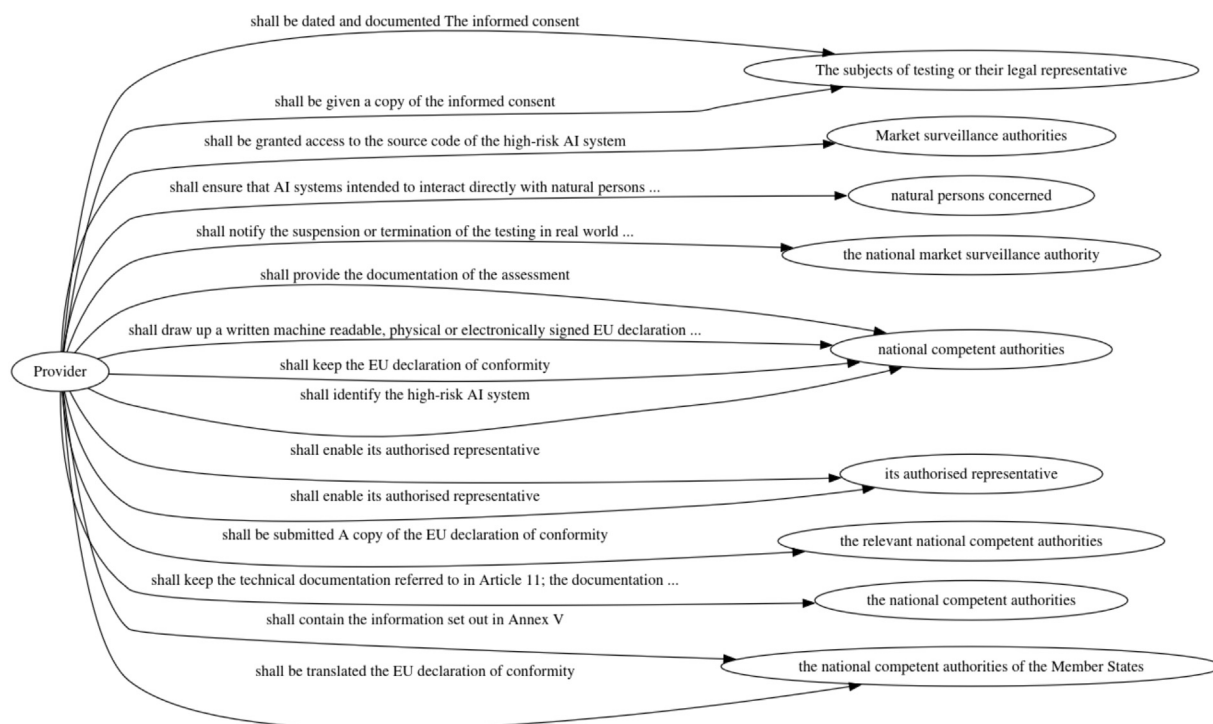


Fig. 3. Sub-graph of the AIA knowledge graph retrieved for the entity “provider”. It shows the associated obligations (predicate and target) and their beneficiaries, as extracted from the AI Act.

text such as the AI Act. In particular, the presented workflow demonstrates the high degree of accuracy that LLMs are capable of when identifying, filtering and analysing deontic obligations. The workflow’s accuracy was evaluated through its application to a set of obligations, yielding an accuracy of 93% in the filtering process and near-perfect accuracy in the classification of obligation types and core elements, such as predicates and addressees.

This approach has the potential to automate the extraction and structuring of obligations, thereby reducing the manual burden typically associated with compliance analysis. The system can expedite the tracing of obligations, encourage more uniform interpretation, and augment the management of compliance risks. Although this study focuses on the AI Act as a timely and significant case study, the methodology itself is generalisable: it relies on linguistic markers and modular processing steps that can be adapted with minimal changes to other regulatory documents and domains. In this sense, the AI Act serves as a proof-of-concept application, while the broader contribution lies in demonstrating a reproducible and transferable workflow. The integration of knowledge graphs further strengthens this generalisability by providing a semantically rich and visually navigable representation of obligations, supporting contextual understanding and query-based exploration across diverse legal domains.

Our experiment also showed some limitations. The system demonstrates deficiencies in its handling of more complex and nuanced elements, such as Specifications and Targets. Moreover, the low KA scores in the Deontic Obligation Analysis show that some of the elements should be more clearly defined in the prompt design as well as in the evaluation procedure. In future research we should consider refining these aspects by introducing more granular categories and enhancing prompt adaptability, possibly also extending the implementation and evaluation to encompass a more extensive set of EU regulations.

Furthermore, while surrounding context has been shown to enhance LLM performance, a more comprehensive integration of cross-referenced legal provisions may also be required to adequately capture the intricacies of legal relationships.

In our contribution, the ontology for deontic obligation representation is only partially developed: its current structure supports the

organisation of entities, actions, and relationships, but remains less mature than the other modules of the workflow. A more comprehensive design and deeper formalisation of the ontology is therefore left to future work, where it may be elaborated as a standalone contribution. This may include also legal definitions, as contained in the Regulation, which could promote consistency in the LLM analysis and further improve the precision of automated classification tasks.

Finally, a future plan is to evaluate alternative LLMs in order to ascertain whether different architectures offer improved performance in the processing of legal text. The employment of a modular approach, which involves the utilisation of multiple specialised LLMs for the execution of distinct subtasks, has the potential to enhance precision and scalability.

**Declaration of competing interest**

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

The research has partially been supported by the “CompuLaw” project, funded by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No. 833647).

**Appendix A. Prompt engineering for deontic obligation filtering**

This section delineates the methodology employed for prompt engineering in the context of the Deontic Obligation Filtering module.

A Chain of Instructions (CoI) [38] approach was adopted, whereby legal experts and computer scientists collaboratively provide the language model with detailed, step-by-step instructions to classify provisions into predefined categories (see Section 4), yielding a structured JSON as output.

The prompt consists of two parts: the *system prompt*, containing the instruction for the task with both expected input and outputs, and the *user prompt*, containing the actual input for the task.

Due to the considerable length of the prompts (approximately 4000 words together), we provide an overview of the key instructions of the system prompt in Listing 3 and of the user prompt in Listing 4.

**Listing 3** System Prompt Overview for Obligation Filtering

```

1 # Classification of potential legal
  obligations
2 You are an expert in analysing regulatory
  texts.
3
4 ## Context
5
6 Your task is to classify each occurrence of
  the term "shall" or "must" or "has/have
  to" into one of the following categories
  based on its use in the text:
7
8 [Presenting the details of Definitions,
  Constitutive statements, Deontic
  Obligations and Prohibitions,
  Entitlements, Authorisations.]
9
10 ## Instructions:
11
12 [A set of instructions on how to analyse and
  extract the required elements and how to
  structure the JSON output.]
13
14 [Few shot examples]
15
16 [Final Remarks]
```

The deontic obligation filtering prompt is composed of two elements. Firstly, there is the *system prompt*, the function of which is to define the task and require the model to justify its classifications. Secondly, there is the *user prompt*, the function of which is to provide the sentence to be analysed. Few-shot examples illustrate candidate obligations paired with their expected classifications.

**Listing 4** User Prompt and Expected Output for the Obligation Filtering Step

```

1 # User prompt
2
3 ## Sentence to analyse
  [Sentence here]
4
5
6 ## Context
  [Paragraph and surrounding paragraphs]
```

The user prompt presents a single sentence along with its surrounding paragraph in order to provide context. This sentence-level focus has been demonstrated to enhance classification accuracy. Citations were excluded from the analysis, as preliminary tests indicated that they exerted minimal influence on the model performance of the obligation filtering step.

## Appendix B. Prompt engineering for obligation analysis

This section delineates the methodology employed for prompt engineering in the context of the Deontic Obligation Analysis step.

Here, a CoI approach was also adopted, whereby legal experts and computer scientists collaboratively provided the language model with detailed, step-by-step instructions to extract, either from the provision, the context, citations or background-knowledge the elements from deontic obligations and prohibitions (see Section 4), yielding a structured JSON as output.

Listing 5 reports an overview of the key instructions in the system prompt, while Listing 6, presents the user prompt, which includes the provisions subject to analysis and the surrounding context and citations.

**Listing 5** System Prompt Overview for Obligation Analysis

```

1 # **Legal Syntactical Analysis of Legal
  Obligations**
2 You are an expert in legal linguistics and
  you are required to make a syntactical
  analysis of legal obligations.
3
4 ## **Obligation of Action vs. Obligation of
  Being**
5 [Definitions and short examples]
6
7 ## **Syntactical Analysis Structure**
8
9 ### **Obligation of Being Elements**
10 [Definitions and short examples]
11
12 ### **Obligation of Action Elements**
13 [Definitions and short examples]
14
15 ## **Extraction Method**
16 [Definitions of "Stated", "Context",
  "Citation", and "Background-Knowledge"]
17
18 ## **Output Structure**
19
20 [Presentation and explanation of the expected
  JSON Structure]
21 [Few shot examples]
22 [Additional guidelines]
23 [Final remarks]
```

The prompt begins with a high-level task definition to ensure clarity. Then, it introduces the distinction between the obligation of being and the obligation of action, supported by their definitions and examples. Next, the obligation extraction framework is presented, detailing the key components as detailed in our theoretical framework: Addressee, Predicate, Target, Specifications, Pre-conditions and Beneficiaries. Each of these elements is interpreted differently depending on the type of deontic obligation (action/being). Note that multiple values can be extracted for each element of the framework (e.g., there may be more than one addressee for the same obligation).

**Listing 6** User Prompt Overview for Obligation Analysis

```

1 ## Sentence to analyse
2 [Content of the sentence to be analysed]
3
4 ## Context
5 [Surrounding text where the sentence appears]
6
7 ## Citations to other paragraphs, sections,
  if any
8 [Content of citations if any]
```

In the *user prompt*, the materials required for the LLM to analyse are presented clearly. This includes the sentence to be analysed, the context, and any relevant citations.

## Data availability

Code, data and prompts are available at [https://github.com/thiagordp/obligation\\_extraction\\_for\\_compliance](https://github.com/thiagordp/obligation_extraction_for_compliance).

## References

- [1] Mahler T. Between risk management and proportionality: The risk-based approach in the EU's Artificial Intelligence Act Proposal. In: *Nordic Yearbook of Law and Informatics*. 2021.
- [2] Kiyavitskaya N, Zeni N, Breaux TD, Antón AI, Cordy JR, Mich L, et al. Automating the extraction of rights and obligations for regulatory compliance. In: *Conceptual modeling-ER 2008: 27th international conference on conceptual modeling*, Barcelona, Spain, October 20-24, 2008. proceedings 27. Springer; 2008, p. 154–68.
- [3] Wyner A, Peters W. On rule extraction from regulations. In: *Legal knowledge and information systems*. IOS Press; 2011, p. 113–22.
- [4] Gao X, Singh MP. Extracting normative relationships from business contracts. In: *Proceedings of the 2014 international conference on autonomous agents and multi-agent systems*. 2014, p. 101–8.
- [5] Neill JO, Buitelaar P, Robin C, Brien LO. Classifying sentential modality in legal language: a use case in financial regulations, acts and directives. In: *Proceedings of the 16th edition of the international conference on artificial intelligence and law*. 2017, p. 159–68.
- [6] Hassani S, Sabetzadeh M, Amyot D, Liao J. Rethinking legal compliance automation: Opportunities with large language models. 2024, arXiv preprint arXiv:2404.14356.
- [7] Kerrigan S, Law KH. Logic-based regulation compliance-assistance. In: *Proceedings of the 9th international conference on artificial intelligence and law*. 2003, p. 126–35.
- [8] Zhou Y-C, Zheng Z, Lin J-R, Lu X-Z. Integrating NLP and context-free grammar for complex rule interpretation towards automated compliance checking. *Comput Ind* 2022;142:103746.
- [9] Jiang J, Aldewereld H, Dignum V, Wang S, Baida Z. Regulatory compliance of business processes. *AI Soc* 2015;30(3):393–402.
- [10] Boella G, Janssen M, Hulstijn J, Humphreys L, Van Der Torre L. Managing legal interpretation in regulatory compliance. In: *Proceedings of the fourteenth international conference on artificial intelligence and law*. 2013, p. 23–32.
- [11] Kiyavitskaya N, Zeni N, Mich L, Cordy JR, Mylopoulos J. Text mining through semi automatic semantic annotation. In: *Practical aspects of knowledge management: 6th international conference, PAKM 2006, Vienna, Austria, November 30-December 1, 2006. proceedings 6*. Springer; 2006, p. 143–54.
- [12] Dragoni M, Villata S, Rizzi W, Governatori G. Combining NLP approaches for rule extraction from legal documents. In: *1st workshop on mining and reasoning with legal texts*. 2016.
- [13] Walzl B, Muhr J, Glaser I, Bonczek G, Scepankova E, Matthes F. Classifying legal norms with active machine learning. In: *Legal knowledge and information systems*. IOS Press; 2017, p. 11–20.
- [14] Liga D, Robaldo L. Fine-tuning GPT-3 for legal rule classification. *Comput Law & Secur Rev* 2023;51:105864.
- [15] Francesconi E, et al. Transfer learning for deontic rule classification: The case study of the GDPR. In: *Legal knowledge and information systems: JURIX 2022: the thirty-fifth annual conference*, vol. 362, IOS Press; 2022, p. 200.
- [16] Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. In: Larochelle H, Ranzato M, Hadsell R, Balcan M, Lin H, editors. *Advances in neural information processing systems*, vol. 33. Curran Associates, Inc.; 2020, p. 1877–901, URL [https://proceedings.neurips.cc/paper\\_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf).
- [17] Hohfeld WN. Fundamental legal conceptions as applied in judicial reasoning. *Yale Law J* 1917;26(8):710–70.
- [18] Kelsen H. *The Pure Theory of Law*. University of California Press; 1967.
- [19] Hart HL. *The Concept of Law*. Oxford University Press; 1961.
- [20] Nguyen H-T, Fungwacharakorn W, Satoh K. Enhancing logical reasoning in large language models to facilitate legal applications. 2023, arXiv preprint arXiv:2311.13095.
- [21] Manas K, Zwicklbauer S, Paschke A. TR2mtl: LLM based framework for metric temporal logic formalization of traffic rules. 2024, arXiv preprint arXiv:2406.05709.
- [22] Sandrelli A, et al. A corpus-based study of deontic modality in English Eurolect. *ESP Across Cult* 2020;(17):113–28.
- [23] European Commission. *English Style Guide: A handbook for authors and translators in the European Commission*. 2016, Updated 2025, p. 59, point 10.27, URL <https://commission.europa.eu/document/download/c45f5b70-2d0e-4da7-b181-b5fe3a16c4bb.en>.
- [24] Bhatia VK. *Vagueness in normative texts*, vol. 23, Peter Lang; 2005.
- [25] Felici A. 'Shall' ambiguities in EU legislative texts. *Comp Legilinguistics* 2012;10(1):51–66.
- [26] Roversi C. In defence of constitutive rules. *Synthese* 2021;199(5):14349–70.
- [27] von Wright GH. *Norm and action*. Routledge and Kegan Paul; 1963.
- [28] Prakken H, Sergot M. *Contrary-to-duty obligations*. *Studia Logica* 1996;57:91–115.
- [29] Fields J, Chovanec K, Madiraju P. A survey of text classification with transformers: How wide? How large? How long? How accurate? How expensive? How safe? *IEEE Access* 2024;12:6518–31. <http://dx.doi.org/10.1109/ACCESS.2024.3349952>.
- [30] Dagdelen J, Dunn A, Lee S, Walker N, Rosen AS, Ceder G, et al. Structured information extraction from scientific text with large language models. *Nat Commun* 2024;15(1). <http://dx.doi.org/10.1038/s41467-024-45563-x>.
- [31] Krippendorff K. Computing krippendorff's alpha-reliability. 2011, URL [https://repository.upenn.edu/asc\\_papers/43](https://repository.upenn.edu/asc_papers/43). [Accessed 19 September 2025].
- [32] Alfieri C, Caroccia F, Inverardi P, et al. AI act and individual rights: A juridical and technical perspective. In: *IAIL@ HHAI*. 2022.
- [33] Almada M, Petit N. The EU AI act: a medley of product safety and fundamental rights? 2023, Robert Schuman Centre for Advanced Studies Research Paper 2023/59.
- [34] Novelli C, Hacker P, Morley J, Trondal J, Floridi L. A robust governance for the AI act: AI office, AI Board, scientific panel, and national authorities. *Eur J Risk Regul* 2024;1–25.
- [35] Varošanec I. On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI. *Int Rev Law, Comput Technol* 2022;36(2):95–117.
- [36] Oduro S, Moss E, Metcalf J. Obligations to assess: Recent trends in AI accountability regulations. *Patterns* 2022;3(11).
- [37] Engler AC, Renda A. Reconciling the AI value chain with the EU's artificial intelligence act. *CEPS*; 2022.
- [38] Hayati SA, Jung T, Boddington T, Kar S, Sethy A, Kim J-K, et al. Chain-of-instructions: Compositional instruction tuning on large language models. 2025, arXiv:2402.11532.