

Unconstitutionality Prevention in Bills Using Hybrid AI

Michele CORAZZA ^a, Pier Francesco BRESCIANI ^a, Generoso LONGO ^a,
Faria FEROOZ ^a, Salvatore SAPIENZA ^a and Monica PALMIRANI ^{a,1}

^a*Department of Legal Studies, CIRSIFID - ALMA AI, University of Bologna*

ORCID ID: Michele Corazza <https://orcid.org/0000-0002-7288-6635>, Pier Francesco
Bresciani <https://orcid.org/0000-0002-6889-1618>, Generoso Longo
<https://orcid.org/0009-0003-2687-5884>, Faria Ferooz
<https://orcid.org/0000-0002-8150-7852>, Salvatore Sapienza
<https://orcid.org/0000-0002-5429-5217>, Monica Palmirani
<https://orcid.org/0000-0002-8557-8084>

Abstract. One of the most important legal assessments during legislative drafting proposals in Parliament is to check whether the bills have some provisions clashing with the Constitutional norms. Sometimes this situation is intentionally created by the proposer of a party to underline inconsistencies or incompleteness in the legal systems according to the reality of the society, but frequently, they are mistakes to avoid. This paper presents a method to calculate a multi-factor index of presumption of unconstitutionality of the provisions of a bill, to inform the proponent concerning a possible constitutional conflict. The paper uses a large dataset of bills, coming from the Chamber of Deputies of Italy, combined with Constitutional Court judgments and abstracts, EU and National legislation (9095 Italian laws, 3710 bills, 2979 judgments from the Constitutional Court). The methodology used is hybrid, considering different technological components, linked in a pipeline: i) embedding similarity; ii) normative references navigation; iii) semantic topic of the bill; iv) legislation cited in the judgments score based on the abrogated.

Keywords. Constitutional law, embedding, Akoma Ntoso, similarity, sentence transformer

1. Introduction and Research Questions

Constitutional review is a cornerstone activity of modern democratic governance aimed at ensuring that legislation aligns with the foundational principles enshrined in a nation's constitution. While *ex post* constitutionality review is usually demanded to specialised high-level courts, it is also carried out *ex ante* (i.e., before the bill is approved) by Parliaments on bills under discussion, to prevent the unconstitutionality of the forthcoming text. Such preliminary constitutionality review is traditionally carried out through legal scrutiny by parliamentary committees and designated departments of civil servants. However, the increasing complexity of legislative proposals and the evolving sociopolitical landscape suggest the possibility of integrating Artificial Intelligence (AI) systems in support of the parliamentary activities for this specific task. From a legal point of view,

¹Corresponding Author: Monica Palmirani, monica.palmirani@unibo.it

this integration is not trivial. Principles of the rule of law and, in particular, constitutional law foundations [1] - e.g., separation of powers, the autonomy of the decision-maker [2], the law-making process - shall be conjugated with technical solutions that ensure accountability, fairness, and explainability of the outputs [3].

In this study, we investigate the feasibility of a novel automated assessment method of the constitutionality of bills while conjugating legal principles and the peculiarities of legal texts. We focus on the Italian legal system, a civil law country in which a Constitutional Court is tasked with an *ex post* assessment, and in particular on the activities of the Chamber of Deputies, one of the two branches of the Parliament, that has the obligation to check the possible unconstitutional provisions.

Given these premises, this study aims to investigate the following research questions:

- **RQ1.** How is it possible to represent the text of a bill with embedding while respecting its structure (paragraphs, subparagraphs, etc.) and navigating normative references point in time? This is a preliminary objective that will allow us to compare the bill with other material (e.g., decisions of the Constitutional Court).
- **RQ2.** How is it possible to fruitfully retrieve relevant decisions from the Italian Constitutional Court, given the text of a bill? This objective allows us to identify the decisions to compare with our bill.
- **RQ3.** How is it possible to compare a bill and the relevant decisions, and then to measure the risk of a bill being unconstitutional, given the bill and the correlated relevant decisions of the Constitutional Court?

To tackle these challenges, this study relies on a hybrid approach which combines an XML-structured representation of the legislative text in Akoma Ntoso and subsymbolic approaches leveraging on this structure [4]. Owing to the novelty and experimental nature of the approach, our results should be interpreted as preliminary, yet they are nonetheless encouraging. The paper is divided as follows: following this introduction, Section 2 introduces the background and the motivation underlying the study; Section 3 describes the datasets, whereas Section 4 details the methodology. Experimental settings are presented in Section 5, while their results are presented in Section 6 and discussed in Section 7. Some conclusive remarks are reported in Section 8.

2. Background and Motivation

In 2024, the 32% of the judgments of the Italian Constitutional Court were about the constitutional illegitimacy of law at the State and regional level (417 decisions, 132 of unconstitutionality) ². It is a problem of democracy, dialogue between institutions, and the division of powers because the Constitutional Court seems to invade the legislative space, especially in those domains with a lack of solid regulation (e.g., end of life). For this reason, the topic is urgent and very sensitive.

However, this task is not very well studied by the scientific research, that has concentrated mostly on prediction, factors classifications, summarization of the decisions.

Predicting courts decision is a classic task of the "AI & Law" domain [5,6]. Relevant studies in the field use NLP techniques in selected domain, and include notorious contributions in the European Court of Human Rights [7,8]. Among them, some studies focus on higher courts tasked with constitutionality assessment (e.g., "Supreme Courts",

²https://www.cortecostituzionale.it/annuario2024/pdf/Dati_quantitativi_e_di_analisi_2024.pdf Annual Report 2024

“High Court”, and so forth). This has been the case for studies focusing on the United States [9], Brasil [10], Germany [11], Singapore [12], and many other countries.

Surveys show a consistent pattern of duality: on the one hand, certain studies focus on case law retrieval from higher courts and have the goal of retrieving pertinent to a set of facts (in most cases) or a given piece of legislation [13]; on the other hand, some studies focus on the constitutionality assessment as a classification task in which the study’s goal is to predict the outcome of the Court’s decision given a set of facts or prior cases [14]. This can happen with or without the case retrieval. Large Language Models (LLM) are showing advances in both fields [15,16,17], despite the chance of hallucinations[18] and technical constraints (e.g., context length).

Our study differs from the existing literature on three grounds:

- it focuses on the Italian legislative system, which has not been subject to particular attention in the domain of unconstitutionality evaluation. This might be due to several reasons, including data retrieval and language complexity;
- it processes bills rather than in-force legislation or sets of facts. While training datasets of the existing literature use applicable legislation, we directly measure the risk of unconstitutionality over bills. The difference lies on the fact that in-force laws are already interpreted by lower courts and constitutional courts have prior interpretations at their disposal. From a legal standpoint, assessing the constitutionality of bills is harder when compared to applicable legislation;
- it uses an XML-enriched version of the legal texts to allow the point-in-time navigation of legal references, rather than plain text. This contributes to increase the proximity between the reasoning performed by the automated system and the one performed by the Constitutional Court.

3. Datasets

In our experiments, we used three distinct datasets from different legal resources: i) Italian legislation; ii) bills; iii) Constitutional Court decisions. Crucially, all these dataset are provided in Akoma Ntoso XML or they were converted by us for our experiments. The three datasets that we used are:

- **Italian legislation in Normattiva**: a dataset of Italian legislative documents extracted from the official website Normattiva containing Italian laws. It is composed of all documents from 2000 to 2025, 9095 in total, including all consolidated versions. Furthermore, we extracted information about the abrogated articles of these documents, which are used in our user interface to show judgments that include normative references to abrogated norms. This parameter is fundamental for informing legal experts regarding the up-to-date status of the legislative information.
- **Chamber of Deputies Bill (PDL)**: a dataset of 3710 legislative bills from the XVIII and XIX (March 2018 to May 2024) legislatures. They were downloaded from the official website of the Italian Chamber of Deputies and converted to Akoma Ntoso XML using an ad-hoc procedure³. In the paper now we mention this dataset simply as CD Italian bills.
- **Italian Constitutional Court Decisions**: our starting point is a dataset containing all the judgments from the Italian Constitutional Court converted to Akoma Ntoso. We first extracted the more recent ones (from 2001 to 2025). Then, the metadata

³<https://gitlab.com/CIRSFID/html2aknPDL>

Result (Italian)	English translation	Annotation
ill. cost. conseg. ex art. 27 legge n. 87/1953	consequential unconstitutionality ex Art. 27 Law No. 87/1953	negative
ill. cost. parziale conseg. ex art. 27 legge n. 87/1953	consequential partial unconstitutionality ex Art. 27 Law No. 87/1953	negative
illegittimità costituzionale	unconstitutionality	negative
illegittimità costituzionale parziale	partial unconstitutionality	negative
manifesta infondatezza	manifest unfoundedness	positive
non fondatezza	unfoundedness	positive

Table 1. The relevant results of the judgments and their annotation

associated with each document was used to extract judgments that contain at least one verdict of unconstitutionality, obtaining 2979 judgments. The full list of verdicts, as they are shown in the metadata, can be found in Table 2.

In addition to selecting the verdicts, it was also crucial to select the verdict results and their relative annotation as either positive (non-unconstitutionality was found) or negative (verdict of unconstitutionality). These result types, as shown in the metadata, are found in Table 4.

4. Methodology

Our approach is to build a meaningful index of unconstitutionality for CD Italian bills according to legal principles and considering the similarity of the bill with the relevant decisions of the Constitutional Court that have already decided unconstitutionality. While an alternative approach could have lead to find similarities between the current bill and legislation already scrutinised by the Court, this would have lead to the inclusion of legislation already assessed and, in turn, some degree of bias towards the prediction. Therefore, we have considered three parameters: the similar domain, the similar legal concepts, and the similar rationale in the legal reasoning.

In order to produce a meaningful measure of the risk of unconstitutionality, the first step is to match bills with the relevant judgments from the Constitutional Court. One might be tempted to use a lexical comparison based on TF-IDF or similar techniques. These approaches, however, rely on the assumption that the lexical overlap between the two documents being compared is significant. In the case of the comparison between bills and constitutional court judgments, these are very different types of documents, which have a different structure and lexicon. For this reason, our approach is based on a sentence transformer [19], a type of model which allows the creation of embeddings for sequences of text, which can then be used to find semantically similar texts. This is typically achieved by using a contrastive learning loss, which forces the vectors for similar sentences to have a high cosine similarity.

While sentence embeddings can be used to compare the semantics of relatively short portions of text, the ability of any model to deal with longer documents is not guaranteed. In fact, even LLMs struggle with documents, recent results show that the effective length of most documents is lower than the maximum number of tokens by at least an order of magnitude [20] and they are effectively able to deal with less than 10k tokens. To produce a meaningful semantic similarity metric between Italian bills and Italian Constitutional Court judgments, we considered the following portions of these documents:

- **CD Italian bills:** by leveraging the Akoma Ntoso format we are able to access the

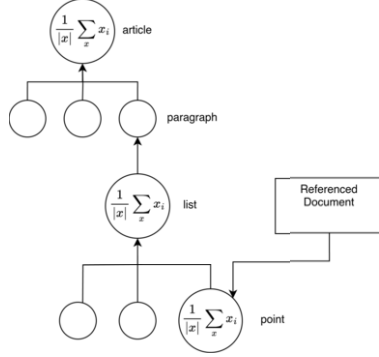


Figure 1. A graphical visualization of a possible application of the tree aggregation strategy. In this case, an article contains three paragraphs, and the last one contains a list of three points, the last point contains a normative reference.

structure of Italian bills and we operate at the granularity of a single article.

- **Chamber of deputies commission:** we use an unsupervised approach to associate each bill with the relevant Italian Chamber of Deputies commission. This is fundamental in order to detect the legal topics of the bill from the legislative competencies of the committee.
- **Constitutional Court decisions:** we select the paragraphs that are associated with the motivation of the judgment, using the corresponding tag of AKN <motivation>.
- **Constitutional Court abstracts:** these are abstracts (massime) which are redacted directly from the Constitutional Court. There can be more than one associated with each decision, as each one may contain more than one interpretation if the constitutionality assessment spans across different provisions.
- **Constitutional Court ratio decidendi:** a series of terms are usually associated by the judges to the decision, remarking the main legal reasons which are associated with each judgment directly from the court.

The discussion of how these different textual sources were used in our approach can be found in the next sections.

4.1. Tree embeddings

Both bills and decisions from the Constitutional Courts, however, have a hierarchical structure (bill has articles, commas, paragraphs, etc. and decisions introduction, background, motivation, decision) and they can contain legal citations. For this reason, we adopt a tree aggregation strategy for our embeddings. For the sake of clarity, we describe this approach as it is applied to articles of CD Italian bills. For the paragraphs of the motivation of the Italian Constitutional Court decisions, we use the same procedure, but we treat the references to other judgments in a slightly altered manner.

Given a sentence transformer model, instead of applying it directly to the text of the entire article, which might be too long for the maximum number of tokens for sentence transformer models, we traverse its Akoma Ntoso XML tree, until we find a leaf node. These leaf nodes were defined manually and they are based on a list of XML elements that are marked as inline in the Akoma Ntoso standard (references, dates, modifications,

etc). Our algorithm proceeds from the article to all its children nodes that do not belong to this list. Then, we need to consider normative references. These references can be punctual (reference to a single article, point, section, etc) or they can be generic references (to an entire document). Formally, given an article a which contains children and references, and a sentence transformer model M , we use the following recursive function:

$$v(a) = \frac{1}{2 + |c(a)|} \left(M(t(a)) + \sum_{i=1}^{|c(a)|} v(c_i(a)) + \frac{1}{|r(a)|} \sum_{h=1}^{|r(a)|} R(r_h(a)) \right) \quad (1)$$

Where:

- $t(a)$ is the text contained in the article which is not a part of any of its children
- $c_i(a)$ is the i -th non inline child of the article
- $c(a)$ is the set of all the non inline children of the article a
- $r_h(a)$ is the h -th normative references which is directly contained in $t(a)$
- $r(a)$ is the set of all normative references which are directly contained in $t(a)$

In order to represent normative references, we define the function R as follows:

$$R(i) = \begin{cases} v(i) & \text{if } i \text{ is a punctual reference} \\ \frac{1}{2}M(T(i)) + v(A_1(i)) & \text{if } i \text{ is a generic normative reference} \\ \frac{1}{2}M(I(i)) + v(P(i)) & \text{otherwise} \end{cases} \quad (2)$$

Where:

- $T(i)$ represents the title of the referenced document
- $A_1(i)$ represent the first article of the referenced document
- $T(i)$ represents the introduction of the referenced judgment
- $P(i)$ represent the motivation of the referenced document

By adopting these recursive functions, we obtain a procedure which creates a vector representation for each article by leveraging its structure, and which also considers the normative references contained in the documents.

4.2. Bill-Decisions Similarities

In order to produce a measure of similarity between a bill and the existing decisions from the Constitutional Court, we use three different similarity measures.

4.2.1. Text Similarity

The first measure is the average similarity between each article of a bill and each paragraph from the <motivation> tag of the decision. Formally, given a bill b and a decision j , the similarity is measured as:

$$s_{ap}(b, j) = \frac{1}{|A(b)||P(j)|} \sum_{i=1}^{|A(b)|} \sum_{h=1}^{|P(j)|} sim(v(A_i(b)), v(P_h(j))) \quad (3)$$

Where:

- sim represents cosine similarity;

- $A(b)$ represents all the articles of bill b , $A_i(b)$ is the i -th article of bill b
- $P(j)$ represents all the paragraphs of the motivation of judgment j , $P_h(j)$ is the h -th paragraph of the motivation of judgment j

4.2.2. Legal Concepts Similarity

Another criterion for the comparison of the judgment and the bill is based on the *ratio decidendi* of the judgment (which qualifies the final result of the decision) and the Italian Chamber Commission assigned to the bill (that qualifies the topics and competencies of the bill). Since most of the bills are never assigned a Commission, we first use an unsupervised approach based on similarity to assign the Commission to each bill b as follows:

$$Z(b) = \operatorname{argmax}_h \left(\frac{1}{|A(b)|} \sum_{i=1}^{|A(b)|} \operatorname{sim}(v(A_i(b)), M(k_h)) \right) \quad (4)$$

Where:

- k represents all the Italian Chamber commissions;
- k_h is the textual description of the h -th commission, as found in the Chamber regulation as provided by its implementation in accordance to art. 22, paragraph 1-bis Lettera circolare sugli ambiti di competenza delle Commissioni permanenti, 16 October 1996;

Once we obtained the Commission assigned to each bill, the similarity is calculated as follows:

$$s_{cz}(b, j) = \operatorname{sim}(M(k_{Z(b)}), M(d(j))) \quad (5)$$

Where $d(j)$ is the ratio decidendi associated with the current judgment j .

4.2.3. Reasoning Similarity

Finally, another similarity criterion is the one involving Judgment abstracts and Articles of the bill. Formally, given a bill b and a judgment j , this similarity can be expressed as:

$$s_{am}(b, j) = \max_h \left(\frac{1}{|A|} \sum_{i=1}^{|A(b)|} \operatorname{sim}(v(A_i(b)), M(m_h(j))) \right) \quad (6)$$

Where $m_h(j)$ represents the h -th abstract associated with judgment j . In this case, since one judgment can be associated with multiple abstracts, we opted to use the maximum similarity between the articles of the bill and each of the abstracts. This is motivated by the possibility of one of the verdicts being more relevant for the bill than the others.

4.3. Unconstitutionality prevention index

In order to show the user a measure of risk for unconstitutionality for a given bill, we first select the 10 most similar judgments from the court using one of the similarity metrics from the previous section. Then, the judgments that contain at least one verdict of unconstitutionality are considered, and the similarity value associated with them is

used to scale their relative importance in the risk formula. Formally:

$$R(b) = \frac{\sum_{h=1}^{10} s_h(b) I_h(b)}{\sum_{i=1}^{10} s_i(b)} \quad (7)$$

Where s_i is the i -th similarity value in the 10 most similar judgment to bill b , while $I_h(b)$ is defined as follows:

$$I_h(b) = \begin{cases} 1 & \text{if the } h\text{-th judgment contains at least an unconstitutionality verdict} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

By adopting this method, we produce a risk value which is influenced by the relative similarity of the judgments to the bill and which is able to provide the stakeholders with a meaningful metric for the risk of unconstitutionality associated with each bill.

Italian	English translation
Giudizio di legittimità costituzionale in via incidentale	Incidental controversies on the constitutional legitimacy of laws
Giudizio di legittimità costituzionale in via principale	Direct controversies on the constitutional legitimacy of laws
Questione incidentale di legittimità costituzionale	Incidental question of constitutional legitimacy

Table 2. The types of judgments which are relevant for our analysis

5. Experimental Settings

From the technical point of view, our goal is to compare the three similarity metrics discussed in Section 4 and to assess the overall feasibility of this approach. We opted to use one of the multilingual models from SentenceBERT, namely “paraphrase-multilingual-mpnet-base-v2” [21]. This model is trained from a monolingual Sentence Transformer model, namely “paraphrase-mpnet-base-v2” which is based on MPNet [22]. The adopted approach is based on the teacher-student paradigm, where the goal is to produce a model where the original and translated versions of the same sentence have a high cosine similarity.

In order to assess the performance of our approach, we tasked an expert in constitutional law to produce a list of 10 bills that involve different topics. Then, the expert performed a preliminary analysis by annotating relevant decisions that could be used to assess the risk of unconstitutionality for each bill. In particular, the focus was the creation of a list of subjects that potentially involve matters of constitutionality, and a list of judgments that are the most important precedents for each subject. Then, the expert manually checked the 10 most similar judgments to each selected bill, as produced by the three similarity metrics. Finally, the 10 most similar judgments were annotated as either “relevant” or “not relevant”.

The selected bills (Table 3) were identified by our Constitutional Law expert on the basis of recent Constitutional Court case-law on the topic. Furthermore, to provide a representative sample of different types of bills, both bills that only modify existing laws and new legislation (which might also involve modifications of existing laws) were considered.

Topic	Bill	Type of legislative intervention
End of life	Bill No. 313/2022	New Act
Employment rights on termination	Bill No. 277/2018	Amending Act
Term limits for elected offices	Bill No. 1644/2024	Amending Act
Regional autonomy	Bill No. 1665/2024	New Act
Protection of animals	Bill No. 1109/2023	New Act
Protection of cultural heritage	Bill No. 1297/2023	New Act
Life without parole	Bill No. 1951/2019	Amending Act
Children's surname	Bill No. 256/2022	Amending Act
Adoption by single persons	Bill No. 467/2018	Amending Act
Criminal penalties for drug offences	Bill No. 71/2022	Amending Act

Table 3. Type of topics chosen for the experiment

Similarity Criterion	Top 10 Recall	Top 5 Recall
Articles-Paragraphs	0	0
Commission-Ratio decidendi	0	0
Articles-Abstract	0.51	0.31

Table 4. Results of the experiments with 10 bills

6. Results Evaluation

To measure the performance of the various similarity criteria, we used recall as a metric, since we have access to the ground truth judgments provided by our expert, as well as those that are within the top 10 and top 5 most similar judgments to a given bill. This choice of 10 and 5 judgments was motivated by a mediation between providing users with as many relevant results as possible, while not overwhelming them with too many, which would make the method less useful. The results of our evaluation are shown in Table 4 and they show the recall of the system when using the top 10 and top 5 most similar judgments, respectively. Since we do not have access of a complete list of true positives, as this would require an extensive and time-consuming exploration of all the constitutional court judgments, we are unable to calculate the precision of the system. In our experiments, the comparison between bill articles and the paragraphs of the motivation proved unsuccessful. This can perhaps be explained by the fact that this comparison involves many different portions of text, which makes it less reliable. Furthermore, the differences in language between a judgment and a bill might also have impacted the performance. The comparison between Chamber of Deputies Commissions and *ratio decidendi* from the Court also proved unsuccessful. This result might be due to the fact that the scope of each Commission is very broad, meaning that some details about the specific bills cannot be represented correctly. Finally, our best results come from the comparison between bill articles and abstracts from the Constitutional Court, which is probably due to the more concise and simpler language used in the abstract when they are compared with the overall judgments. Furthermore, since each verdict is associated with its own abstract, this further enhances the granularity of the approach, allowing the embeddings to be more accurate.

7. Discussion

In order to deepen our understanding of the shortcomings of our approach, we conducted a post hoc legal analysis of the results. Firstly, we observed that in two cases (Term Limits for Elected Offices and Criminal Penalties for Drug Offences), our system was unable to retrieve any relevant judgments. From a legal perspective, both bills made only very

specific modifications to highly technical existing legislation. Specifically, the bill on term limits for elected offices consisted of a single substantive article and contained only one explicit reference to the area of legislation affected by the modification. This suggests that our methodology performs better when the theme of the legislation is repeatedly expressed within the legislative texts, whereas it tends to struggle when the context must be inferred primarily through normative references. Secondly, in the case of the bill on children's surnames, the system was able to retrieve all the relevant case-law annotated by our expert. This bill, while modificative in nature, amends the Civil Code, which is a particularly structured type of legislation, with a set of topics that are generally mutually related and with semantically informative article titles. Thirdly, when comparing true positives and false negatives, it can be observed that the approach struggles to identify the most recent relevant judgments. In many cases (e.g., end-of-life or employment rights on termination), the system identified some relevant case-law but missed the most recent judgments on the matter. This limit is legally significant because of the evolving nature of constitutional case-law, which requires constitutionality analyses to begin with the judgments that best reflect the approach of the sitting court. Finally, the system suggested judgments that were not annotated *ex ante* by our expert but were considered relevant in the *post hoc* analysis. This highlights that very complex, time-consuming, and open-ended legal tasks, such as researching the relevant case-law to assess the constitutionality of bills, can benefit from assistive AI-based applications, which are capable of processing large volumes of information in a short amount of time.

8. Conclusive Remarks

This paper demonstrates an innovative hybrid pipeline to calculate an index of unconstitutionality of the bills of the Chamber of Deputies of Italy. We have fostered the AKN-XML annotation of the heterogeneous legal sources repository, using embeddings to prepare vectors for similarity comparison. We have also combined different levels of similarity to grasp different aspects of the legal principles in the Constitutional Law to support the unconstitutionality analysis. The results are promising, but we intend to add more levels of legal analysis in the future. We aim to improve our work, including Question/Answering knowledge graphs coming from the legal knowledge of domain. Secondly, we intend to annotate some *ratio decidendi* with an ontology in order to improve the embeddings with RAG based on supervised legal knowledge. Thirdly, we will start an in-place evaluation made by legal experts of the Chamber of Deputies to capture more reasoning models to inject in our system. Another improvement could be a time-sensitive modeling of the unconstitutionality index, which would consider crucial temporal aspects such as the weight of abrogated norms. Finally, we intend to further improve and validate our results by including articles of the Constitution in the analysis.

Acknowledgments

This project is conducted with the support of the European Commission funds within ERC HyperModeLex. Grant agreement ID: 101055185. This project is co-funded by the European Union - NextGenerationEU under the National Recovery and Resilience Plan (PNRR) - Mission 4 Education and research - Component 2 From research to business - Investment 1.1 Notice PRIN 2022 (DD N. 104 del 02/02/2022), title Smart Legal Order in DigiTal Society (SLOTS), proposal code 2022LRL2C2 - CUP J53D23005610006.

References

- [1] Pier Francesco Bresciani MP. Constitutional Opportunities and Risks of AI in the law-making process. *Federalismie*. 2025;2/2024.
- [2] Pier Francesco Bresciani MP. AI Act and National Parliaments: How the EU is Regulating Democracy in the Digital Era. *Ragion pratica, Rivista semestrale*. 2025;1/2025:249-68.
- [3] Palmirani M. A smart legal order for the digital era: a hybrid AI and dialogic model. *Ragion pratica*. 2022;(2):633-55.
- [4] Palmirani M, Sapienza S, Ashley K, et al. A hybrid artificial intelligence methodology for legal analysis. *Biolaw Journal*. 2024;3:389-409.
- [5] Ashley KD. *Artificial intelligence and legal analytics: new tools for law practice in the digital age*. Cambridge University Press; 2017.
- [6] Ashley KD. A brief history of the changing roles of case prediction in AI and law. *Law Context: A Socio-Legal J*. 2019;36:93.
- [7] Tsarapatsanis D, Aletras N, Preoțiuc-Pietro D, Lampos V. Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective. *PeerJ Computer Science*. 2016 Oct.
- [8] Medvedeva M, Vols M, Wieling M. Using machine learning to predict decisions of the European Court of Human Rights. *Artificial Intelligence and Law*. 2020;28(2):237-66.
- [9] Katz DM, Bommarito II MJ, Blackman J. Predicting the behavior of the supreme court of the united states: A general approach. *arXiv preprint arXiv:14076333*. 2014.
- [10] Bertalan VGF, Ruiz EES. Predicting Judicial Outcomes in the Brazilian Legal System Using Textual Features. In: *DHandNLP@ PROPOR*; 2020. p. 22-32.
- [11] Waltl B, Bonczek G, Scepankova E, Landthaler J, Matthes F. Predicting the outcome of appeal decisions in germany's tax law. In: *International conference on electronic participation*. Springer; 2017. p. 89-99.
- [12] SOH JTH, CHAI IE. Legal area classification: A comparative study of text classifiers on Singapore Supreme Court judgments.(2019). In: *Proceedings of the Natural Legal Language Processing Workshop 2019*, Minneapolis, Minesotta, June. vol. 7;. p. 67-77.
- [13] Feng Y, Li C, Ng V. Legal case retrieval: A survey of the state of the art. In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*; 2024. p. 6472-85.
- [14] Alcántara Francia OA, Nunez-del Prado M, Alatrasta-Salas H. Survey of text mining techniques applied to judicial decisions prediction. *Applied Sciences*. 2022;12(20):10200.
- [15] Kim C, Lee J, Hwang W. LegalSearchLM: Rethinking Legal Case Retrieval as Legal Elements Generation. *arXiv preprint arXiv:250523832*. 2025.
- [16] Deng C, Mao K, Dou Z. Learning interpretable legal case retrieval via knowledge-guided case reformulation. *arXiv preprint arXiv:240619760*. 2024.
- [17] Bertalan VGF, Ruiz EES. Using attention methods to predict judicial outcomes. *Artificial Intelligence and Law*. 2024;32(1):87-115.
- [18] Magesh V, Surani F, Dahl M, Suzgun M, Manning CD, Ho DE. Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. *Journal of Empirical Legal Studies*. 2025;22(2):216-42.
- [19] Reimers N, Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics; 2019. Available from: <https://arxiv.org/abs/1908.10084>.
- [20] Modarressi A, Deilamsalehy H, Dernoncourt F, Bui T, Rossi RA, Yoon S, et al. NoLiMa: Long-Context Evaluation Beyond Literal Matching. In: *Forty-second International Conference on Machine Learning*; 2025. Available from: <https://openreview.net/forum?id=00shX1hiSa>.
- [21] Reimers N, Gurevych I. Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics; 2020. Available from: <https://arxiv.org/abs/2004.09813>.
- [22] Song K, Tan X, Qin T, Lu J, Liu TY. Mpnnet: Masked and permuted pre-training for language understanding. *Advances in neural information processing systems*. 2020;33:16857-67.