

Multilingual Legal Information Retrieval System for Mapping Recitals and Normative Provisions

Rohan NANDA ^{a,b,1}, Llio HUMPHREYS ^c, Lorenzo GROSSIO ^d and Adebayo KOLAWOLE JOHN ^e

^aMaastricht Law and Tech Lab, Faculty of Law, Maastricht University, Netherlands

^bInstitute of Data Science, Maastricht University, Netherlands

^cComputer Science Department, University of Turin, Italy

^dLaw Department, University of Turin, Italy

^eBusinesspoint Intelligence Solutions Ltd, Ireland

Abstract. This paper presents a multilingual legal information retrieval system for mapping recitals to articles in European Union (EU) directives and normative provisions in national legislation. Such a system could be useful for purposive interpretation of norms. A previous work on mapping recitals and normative provisions was limited to EU legislation in English and only one lexical text similarity technique. In this paper, we develop state-of-the-art text similarity models to investigate the interplay between directive recitals, directive (sub-)articles and provisions of national implementing measures (NIMs) on a multilingual corpus (from Ireland, Italy and Luxembourg). Our results indicate that directive recitals do not have a direct influence on NIM provisions, but they sometimes contain additional information that is not present in the transposed directive sub-article, and can therefore facilitate purposive interpretation.

Keywords. legal information retrieval, recitals, European legislation, interpretation

1. Introduction

It is well known in the AI & Law community that norms require legal interpretation: ‘It is clear that these documents are not themselves the law from the fact, that we must first interpret statutes and cases to get at the law which they represent, and from the fact that reasonable persons can disagree as to just what the law is, although there is rarely disagreement as to what, words make up the statute or case in question.’ [5], page 2. Canons of interpretation have been used in Civil and Common Law countries, while the European Court of Justice [9] recommends resolving ambiguous, imprecise or incomplete norms with purposive interpretation (i.e. taking account of the purpose of the norm). Therefore, it is important to note the holistic character of the law in that the meaning of normative provisions often emerges from a wider legislative corpus.

¹Corresponding Author: Rohan Nanda, Maastricht University, Bouillonstraat 3, 6211 LH Maastricht; E-mail:r.nanda@maastrichtuniversity.nl

This paper is concerned with ‘hidden’ links between norms in EU directives and norms in the legislation that transpose them into national law, known as national implementing measures (NIMs). Hidden links are implicit links which are not explicitly referred to within the text of the normative provision via long- or short-form citation. Conceptually similar is one such type of link, but there are others, such as Constitutive, Motivation, Impact etc [3]. Moreover, some of the text in the recitals of the preambles of directives are remarkably similar to some (sub-)articles in the same directive, and are also not made explicit. Recitals can provide additional information and citations to justify the norms in the directive. Den Heijer et al. [1] found that the use of recitals often do not correspond to their stated objectives in official drafting rules [2], and are more significant than commonly appreciated. Due to space constraints, the readers may refer to [6] and [3]. Incidentally, the equivalent to recitals in the countries we looked at consist of generic procedural references with no reference to specific subject-matter of the NIMs. That may not be the case in other Member States and maybe the subject of future work.

There are many different kinds of possible relationships between legal provisions (see [15] and [3]). Previous work on mapping recitals and normative provisions was limited to EU legislation in English and utilized only one text similarity technique [6]. In this paper, we propose, develop and validate a multilingual legal information retrieval system for mapping conceptually similar directive recitals, directive (sub-)articles and NIM provisions. We develop state-of-the-art syntactic and semantic text similarity models to identify conceptually similar norms. The multilingual information retrieval system was validated by evaluating the text similarity techniques on the gold standard mappings (between norms) over a multilingual parallel corpus of 5 directives and their corresponding NIMs from Luxembourg, Ireland and Italy. Our research questions are as follows:

- RQ1) How are directive recitals related to the provisions of the National Implementing Measures (NIMs)?
- RQ2) Which automated text similarity techniques are best able to capture conceptually similar directive recitals, directive (sub-)articles and NIM provisions?
- RQ3) The NIMs of which Member State are most/least semantically correlated with recitals?

2. Methodology

Figure 1 presents the overall workflow of our methodology.

2.1. Corpus generation

A multilingual parallel corpus of 43 directives and their corresponding NIMs from Luxembourg, Ireland and Italy (15,400 norms) was presented in [12]. This corpus only contains mappings between directive (sub-)articles and NIM provisions. It does not contain mappings between directive recitals and directive (sub-)articles, and between directive recitals and NIM provisions. Since the preparation of such fine-grained mappings for the entire multilingual corpus of 43 directives is a highly time-consuming and expensive process, we selected 5 directives and their corresponding NIMs from this corpus. The list of selected directives and their corresponding NIMs is presented in Table 1, and can be found on eur-lex.europa.eu by searching for the CELEX numbers in that table. The title

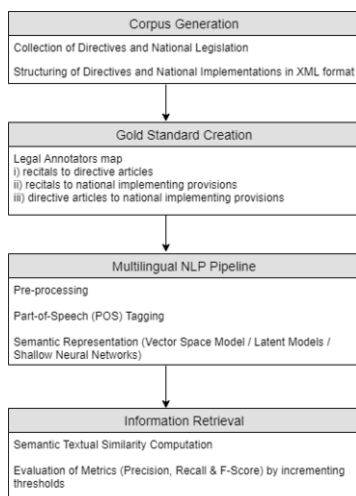


Figure 1. Overall workflow of the multilingual legal information retrieval system

and references of the NIMs for a particular directive can also be obtained on the EUR-Lex portal. This information can be used to obtain the full-text version on the websites of the national legislation portal. The directive recitals, directive (sub-)articles and NIM provisions were all stored in separate XML files for each directive and language.

Table 1. The CELEX numbers of directives and NIMs in the corpus

Directive	NIM (Ireland)	NIM (Luxembourg)	NIM (Italy)
32003L0010	72003L0010IRL_133619	72003L0010LUX_142437	72003L0010ITA_132468
32002L0044	72002L0044IRL_133618	72002L0044LUX_142436	72002L0044ITA_124474
32001L0024	72001L0024IRL_180124 72001L0024IRL_28393	72001L0024LUX_114418	72001L0024ITA_30729
31999L0092	71999L0092IRL_111679	71999L0092LUX_120249	71999L0092ITA_111680
32001L0113	72001L0113IRL_116060	72001L0113LUX_116062	72001L0113ITA_116061

2.2. Gold standard creation

The gold standard corpus was prepared by a researcher in legal informatics, and then checked by a legal expert. The following mappings were prepared for each language: i) directive (sub-)articles to NIM provisions; ii) directive recitals to directive (sub-)articles; and iii) directive recitals to NIM provisions. Mappings were assigned whenever there was content in the two norms that used similar or different wording to express more or less the same content. It was not deemed necessary that the whole text of the norms should be conceptually similar, only a part of it. This is because the similar parts help identify related norms, while the non-similar part should add further information and thus help in interpretation.

2.3. Multilingual NLP pipeline for mapping norms

We developed a multilingual NLP pipeline for mapping directive recitals, directive (sub-)articles and NIM provisions (in this section, described as norms). Using spaCy

(<https://spacy.io/>) tokenizers, we segmented the legal norms into sentences and words, and converted the tokens into lowercase. We removed common noisy words (using spaCy's default list of stopwords) as well as punctuation. We did not select words to retain based on their part-of-speech tag because in the case of short text similarity models (as we are comparing legal norms instead of documents), we need to utilize all the available linguistic features to achieve an acceptable magnitude of text similarity [12]. The text representation phase then encodes the linguistic (syntactic and semantic) features.

For syntactic text representation tests, we utilized: Term Frequency-Inverse Document Frequency (TF-IDF), Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA) and Unifying Similarity Measure (USM)[11] - which combines cosine similarity, N-gram similarity, and approximate string matching, weighted with an arithmetic mean. A dimension size of 50 was chosen for the LSA as it yielded the best performance. The number of topics for the LDA was set to 500 (as it achieved best performance). We used a gram-size of 4 for N-gram similarity in the USM. For our semantic text representation tests, we utilized: FastText (<https://fasttext.cc/>), the Paragraph Vector model [4], and DistilBERT, a lighter model of BERT (Bidirectional Encoder Representations from Transformers) [14]. We used two versions of FastText embeddings: 1)FT-Legal: trained on the complete multilingual parallel corpus including 4,300 directives in English, French and Italian and 27,365 NIMs from Ireland, 14,365 from Luxembourg and 16,233 from Italy, and 2) FT-Generic (<https://fasttext.cc/docs/en/crawl-vectors.html>): pre-trained on Common Crawl and Wikipedia. We used the word-average method, which divides the sum of the word embeddings in a legal norm by the norm length. The embedding dimension size was set to 128. The default hyperparameters were: context window: 5, number of negative samples: 5 and learning rate: 0.1. The Paragraph Vector model [4] was trained on the same multilingual parallel corpus as FastText and used the same embedding dimension size of 128. For the pre-trained DistilBERT embeddings, we used the spaCy-sentence-transformers (<https://spacy.io/universe/project/spacy-sentence-bert>) library to obtain fixed-length (768 dimensions) legal norm vectors.

2.4. Information retrieval

After obtaining the legal norm vectors (see Section 2.3), we compute a cosine similarity measure between them. For instance, to find the most similar NIM provision for a particular directive recital *RI* (case 1 in the above list), a cosine similarity score is computed between the directive recital vector of *RI* and the provisions of the relevant NIM. The NIM provision vectors with a cosine similarity value greater than or equal to the threshold value are retrieved. Each similarity measure is evaluated by comparing the retrieved legal norms with gold standard mappings. Evaluation metrics recall, precision and F-Score are computed for the three types of mappings: directive (sub-)articles to NIM provisions, directive recitals to directive (sub-)articles, and directive recitals to NIM provisions. Evaluation metrics are recorded by incrementing threshold values from 0 to 1 (the increment interval is set to 0.01). The threshold which yields the best F-Score is chosen.

3. Results and analysis

In this section, we present the results of the evaluation of different text similarity measures. The macro-average precision, recall and F-Score are computed for each mapping

type across all three languages. For instance, to map recitals with the NIM provisions of Luxembourg (written in French), the European directive in French was utilized. We discuss the results of each mapping type in the following subsections.

3.1. Mappings between directive recitals and NIM provisions

Figure 2 presents the macro-average precision, recall and F-Score metrics of various text similarity techniques over multilingual mappings between directive recitals and NIM provisions. We observe that the Luxembourg mappings achieved a higher F-Score than the Italian and the English (Ireland) mappings for each similarity measure. This is because of the presence of more common words and phrases between the French directive recitals and the Luxembourg NIM provisions. We also observed that in case of Irish English-language legislation, the precision is much lower than for the French-language (Luxembourg) and Italian legislation. On the other hand, the recall was the highest in the English language. This is because of a higher number of NIM provisions in Ireland (as shown in Table 2) compared to Luxembourg and Italy.

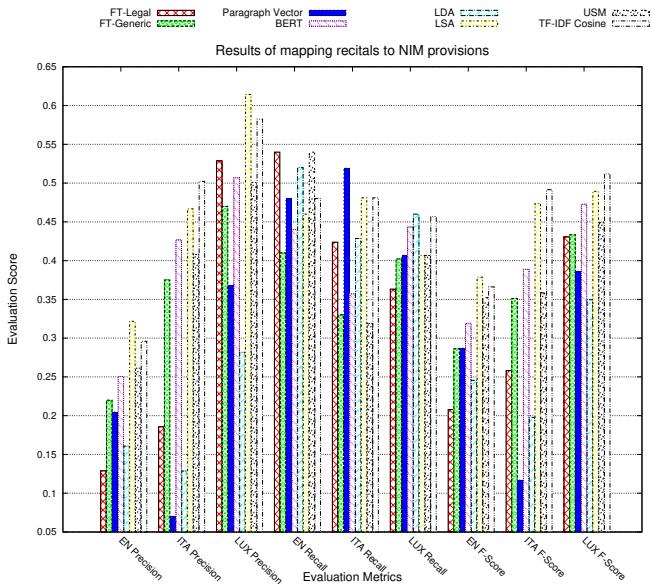


Figure 2. Comparison of the semantic textual similarity techniques for mapping directive recitals and NIM provisions

Due to a high number of provisions in the Irish legislation, the directive recitals shared a decent magnitude of similarity to many unrelated NIM provisions that were not included in the gold standard mapping. This resulted in a higher number of false positives, which led to a low precision score. TF-IDF Cosine and LSA text similarity techniques outperformed other techniques in terms of F-Score in all three languages. This shows that a large number of mappings between directive recitals and NIM provisions

can be identified by weighting important terms through TF-IDF and LSA transform. The performance of semantic text similarity models like FastText (FT-Legal and FT-Generic), Paragraph Vector and BERT was comparatively poorer. However, it is important to note that the performance of BERT model was superior to both FastText and Paragraph Vector. We also observed that the performance of the pre-trained FT-Generic was slightly superior to the domain-specific FT-Legal embeddings.

The best overall macro-average F-Score values are 0.5119, 0.4914 and 0.3786 for mappings between directive recitals and the NIM provisions of Luxembourg, Italy and Ireland respectively. These results indicate that the majority of directive recitals do not share a high degree of semantic similarity with the NIM provisions.

Table 2. The number of provisions in the NIM corpus of each piece of legislation

Ireland NIMs	Luxembourg NIMs	Italian NIMs
269	194	146

3.2. Mappings between directive (sub-)articles and NIM provisions

Figure 3 presents the macro-average precision, recall and F-Score for the mappings between directive (sub-)articles and NIM provisions. These results indicate that the Luxembourg directive mappings consistently achieved a higher recall, precision and F-Score than Italian and English ones for all the similarity measures. This is consistent with the research presented in [12]. The best overall macro-average F-Score values are 0.8243, 0.7276 and 0.6712 for mappings between directive (sub-)articles and the NIM provisions of Luxembourg, Italy and Ireland respectively. The best performance was achieved by TF-IDF Cosine similarity measure in all three languages. Mappings between directive (sub-)articles achieved a much higher F-Score compared to mappings between directive recitals and NIM provisions in all three languages. This is quite intuitive because directive (sub-)articles are supposed to be transposed into the national legislation of Member States. There is no obligation to transpose directive recitals into NIM provisions

3.3. Mappings between directive recitals and directive (sub-)articles

Figure 4 presents macro-average precision, recall and F-Score metrics for mappings between directive recitals and directive (sub-)articles by the best performing measure, TF-IDF Cosine. Gold standard mappings for directive recitals and directive (sub-)articles are the same for all three languages because directives have the same structure and content in all EU languages. The similar F-Scores in this case signify that these mappings are not influenced by language differences. Further, it indicates that these mappings can be identified with the same F-Score in different languages. This result also validates our text similarity and gold standard mapping approach. The minor differences in the F-Score are due to the different NLP pipeline models used for tokenization and splitting sentences.

3.4. Discussion

The results from section 3.1 and 3.2 indicate that the French language mappings had the best F-Scores. Table 3 shows an example of a directive recital / directive sub-article / NIM

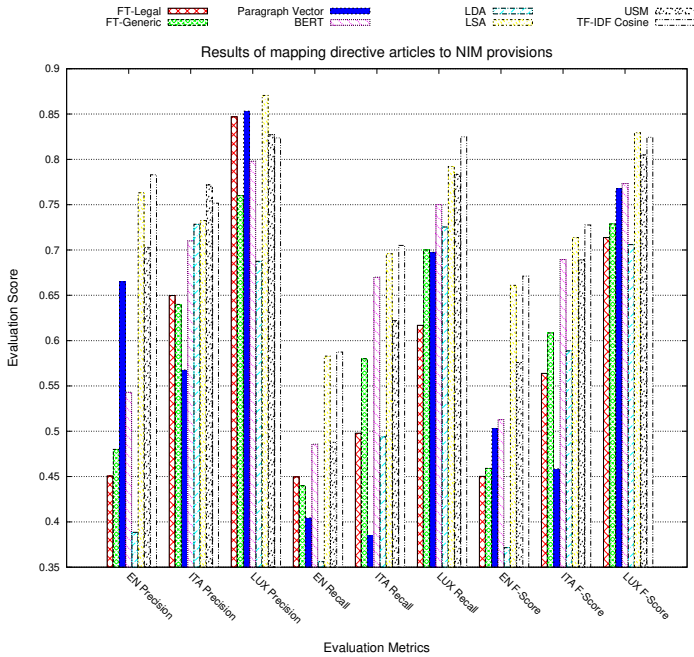


Figure 3. Comparison of the semantic textual similarity techniques for mapping directive (sub-)articles and NIM provisions

Table 3. Similar directive recital, directive sub-article and NIM provision

<p>32002L0044 Directive Recital 8: Pour les secteurs de la navigation maritime et aérienne, dans l'état actuel de la technique, il n'est pas possible de respecter, dans tous les cas, les valeurs limites d'exposition relatives aux vibrations transmises à l'ensemble du corps. Il y a donc lieu de prévoir des possibilités de dérogations dûment justifiées.</p>
<p>32002L0044 Directive Article 10.1: Dans le respect des principes généraux de la protection de la sécurité et de la santé des travailleurs, les États membres peuvent, pour les secteurs de la navigation maritime et aérienne, dans des circonstances dûment justifiées, déroger à l'article 5, paragraphe 3, en ce qui concerne les vibrations transmises à l'ensemble du corps, lorsque, compte tenu de l'état de la technique et des caractéristiques spécifiques des lieux de travail, il n'est pas possible de respecter la valeur limite d'exposition malgré la mise en œuvre de mesures techniques et/ou organisationnelles.</p>
<p>72002L0044LUX.142436 NIM Provision 9.1: Le ministre ayant le travail dans ses attributions peut donner une dérogation à l'article 5, paragraphe 3, dans le respect des principes généraux de la protection de la sécurité et de la santé des travailleurs, pour les secteurs de la navigation maritime et aérienne, dans des circonstances dûment justifiées, en ce qui concerne les vibrations transmises à l'ensemble du corps, lorsque, compte tenu de l'état de la technique et des caractéristiques spécifiques des lieux de travail, il n'est pas possible de respecter la valeur limite d'exposition malgré la mise en œuvre de mesures techniques et/ou organisationnelles.</p>

provision triplet from the French corpus (common texts highlighted). The best performing measure, TF-IDF Cosine was able to identify all three mappings.

There are some cases where the directive recital and NIM provision that are mapped to a directive (sub-)article do not share similar content. An example from the English corpus is shown in table 4 (the similar content between directive recital 9 and directive article 8 are highlighted in yellow, while the similar content between directive article 8 and NIM provision 5.3 are highlighted in green). TF-IDF Cosine identified the link between recital 9 and article 8 (of the directive), and also between directive article 8 and provision 5.3 of the NIM. However, due to the lack of semantic overlap between directive recital 9 and NIM provision 5.3, no similarity link was identified by TF-IDF Cosine. We did not find evidence of direct influence of directive recitals on NIM provisions, which

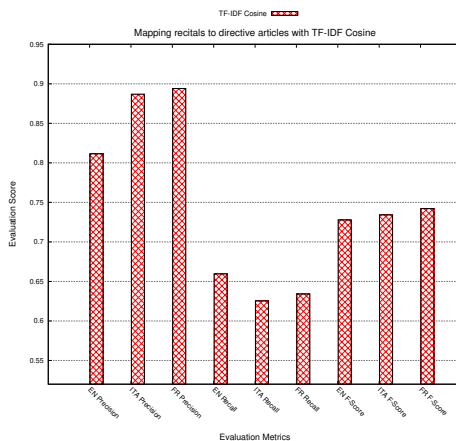


Figure 4. Evaluation metrics for mapping directive recitals to directive articles with TF-IDF Cosine (best performing text similarity measure)

Table 4. Directive article sharing different similar content with directive recital and NIM provision

31999L0092 Directive Recital 9 The establishment of a coherent strategy for the prevention of explosions requires that organisational measures complement the technical measures taken at the workplace; Directive 89/391/EEC requires the employer to be in possession of an assessment of the risks to workers' health and safety at work; this requirement is to be regarded as being specified by this Directive in that it provides that the employer is to draw up an explosion protection document, or set of documents, which satisfies the minimum requirements laid down in this Directive and is to keep it up to date; the explosion protection document includes the identification of the hazards, the evaluation of risks and the definition of the specific measures to be taken to safeguard the health and safety of workers at risk from explosive atmospheres, in accordance with Article 9 of Directive 89/391/EEC; the explosion protection document may be part of the assessment of the risks to health and safety at work required by Article 9 of Directive 89/391/EEC.
31999L0092 Directive Article 8 In carrying out the obligations laid down in Article 4, the employer shall ensure that a document, hereinafter referred to as the "explosion protection document", is drawn up and kept up to date. The explosion protection document shall demonstrate in particular: that the explosion risks have been determined and assessed, that adequate measures will be taken to attain the aims of this Directive, those places which have been classified into zones in accordance with Annex I, those places where the minimum requirements set out in Annex II will apply, that the workplace and work equipment, including warning devices, are designed, operated and maintained with due regard for safety, that in accordance with Council Directive 89/655/EEC(10), arrangements have been made for the safe use of work equipment. The explosion protection document shall be drawn up prior to the commencement of work and be revised when the workplace, work equipment or organisation of the work undergoes significant changes, extensions or conversions. The employer may combine existing explosion risk assessments, documents or other equivalent reports produced under other Community acts.
71999L0092 IRL 111679 NIM Provision 5.3 The risk assessment shall be reviewed by the employer regularly so as to keep it up to date and particularly if— there is reason to suspect that the risk assessment is no longer valid; or there has been a significant change in the matters to which the risk assessment relates including when the workplace, work processes, or organisation of the work undergoes significant changes, extensions or conversions; and where, as a result of the review, changes to the risk assessment are required, those changes shall be made.

would have been evidenced for any triplets by text present in the directive recital and NIM provisions that is absent from the directive article. However, we did find examples of directive recitals containing additional information to related directive (sub-)articles, which can aid purposive interpretation - see table 5 (additional information is in bold).

Table 5. Similar directive recital, directive article and NIM provision, with additional information in the directive recital

32003L0010 Directive Recital 7: Come secondo passo, si ritiene opportuno introdurre misure di protezione dei lavoratori contro i rischi derivanti dal rumore a causa dei suoi effetti sulla salute e sulla sicurezza dei lavoratori, in particolare per quanto riguarda i danni all'udito. Tali misure mirano non solo ad assicurare la salute e la sicurezza di ciascun lavoratore considerato individualmente, ma anche a creare per tutti i lavoratori della Comunità una piattaforma minima di protezione che eviti possibili distorsioni di concorrenza.
32003L0010 Directive Article 1: La presente direttiva, che è la diciassettesima direttiva particolare a norma dell'articolo 16, paragrafo 1, della direttiva 89/391/CEE, stabilisce prescrizioni minime di protezione dei lavoratori contro i rischi per la loro salute e sicurezza che derivano, o possono derivare, dall'esposizione al rumore e, segnatamente, contro il rischio per l'udito.
72003L0010 ITA_132468 NIM Provision 49_bis: Il presente titolo determina i requisiti minimi per la protezione dei lavoratori contro i rischi per la salute e la sicurezza derivanti dall'esposizione al rumore durante il lavoro e in particolare per l'udito.

4. Related work

Most work on links between norms in legislation e.g. [13][16] focus on the discovery and classification of explicit citations. Amantea et al. [3] proposed a model for classifying different kinds of implicit links between directive recitals and directive (sub-)articles including Conceptually Similar, Constitutive, Motivation and Impact. The authors suggested that different kinds of algorithms are required to identify each kind of link, but none were tested. Humphreys et al. [6] mapped recitals to legal articles (but not sub-articles) in EU legislation based on conceptual similarity. Norms were modeled as TF-IDF vectors and similarity was computed based on Cosine Similarity. The system achieved a high recall but low precision. The high accuracy achieved was due to the unbalanced dataset, with a great number of true negatives. Nanda et al. [11][12] investigated automated mapping of directive (sub-)articles to NIM provisions using a variety of similarity algorithms suited for short text including matching common words, common sequences of words and approximate string matching. The relevance of directive recitals was not considered for this work. The work of Lau [8] concerns finding similar provisions in different legislation in the US. A list of the most similar pairs of provisions are produced based on the similarity of parsed norms as well as associated features including legislative definitions and glossaries from reference books. Kumar et al. [7] also used a range of factors to find similar judgments from the Supreme Court of India including headnote, citation and case citation. The most important features were legal terms and citations. Legal-term cosine similarity performed better than all-term cosine similarity.

5. Conclusions and future work

This paper was concerned with ‘hidden’ links between norms in EU directives and national implementing measures. Automated identification of such links could facilitate purposive interpretation and monitoring of implementation. We focussed on identifying conceptually similar norms, and evaluated the performance of suitable text similarity techniques. Since the preparation of fine-grained provision mappings is time-consuming and expensive, we limited our experiments to five directives and their corresponding NIMs. Out of many text similarity techniques, the best performing model was TF-IDF Cosine Similarity. This is consistent with other research in the legal domain [12]. The semantic text representation methods in particular performed adequately for Luxembourg but poorly for the other countries. We found conceptually similar directive recitals, directive (sub-)articles and NIM provisions in all five directives and related NIMs in Ireland, Luxembourg and Italy. However, there was less similarity between directive recitals and NIM provisions, since NIMs are meant to transpose (sub-)articles, not recitals. We did not find evidence of direct influence of directive recitals on NIM provisions. However, we did find directive recitals that contain additional information that can facilitate purposive reasoning. The degree of similarity between directive recitals and NIM provisions varied according to country in exactly the same way as for directive (sub-articles), with NIMs from Luxembourg bearing the highest similarity, and NIMs from Ireland the lowest. The similar F-score in mappings between directive recitals and directive (sub-)articles for different language versions of the same directive shows that our approach is generally sound. However, one reason for the low F-score for mappings to Irish NIMs could be

imprecision and inconsistency in EU English legal language due to the legal drafting being carried out by non-native English speakers who are unfamiliar with Common Law systems and terminology [17]. Our future work will investigate whether mapping equivalent terms in EU directives and NIMs (through an ontology [10]) can improve the performance of the text similarity system and facilitate the detection of conceptually similar normative provisions.

References

- [1] Maarten den Heijer, Teun van Os van den Abeelen, and Antanina Maslyka. On the use and misuse of recitals in European Union law. *Amsterdam Law School Research Paper*, (2019-31), 2019.
- [2] European Commission. Joint Practical Guide of the European Parliament, the Council and the Commission for Persons Involved in the Drafting of European Union Legislation. 2016.
- [3] Ilaria Angela Amantea, Luigi Di Caro, Llio Humphreys, Rohan Nanda, and Emilio Sulis. Modelling norm types and their inter-relationships in EU directives. In *ASAIL@ ICAIL*, 2019.
- [4] Andrew M Dai, Christopher Olah, and Quoc V Le. Document embedding with paragraph vectors. *arXiv preprint arXiv:1507.07998*, 2015.
- [5] Thomas F Gordon. The role of the exceptions in models of the law. *Formalisierung im Recht und Ansätze juristischer Expertensysteme*, pages 52–59, 1986.
- [6] Llio Humphreys, Cristiana Santos, Luigi Di Caro, Guido Boella, Leon Van Der Torre, and Livio Robaldo. Mapping recitals to normative provisions in EU legislation to assist legal interpretation. In *JURIX*, pages 41–49, 2015.
- [7] Sushanta Kumar, P Krishna Reddy, V Balakista Reddy, and Aditya Singh. Similarity analysis of legal judgments. In *Proceedings of the Fourth Annual ACM Bangalore Conference*, pages 1–4, 2011.
- [8] Gloria T Lau, Kincho H Law, and Gio Wiederhold. Similarity analysis on government regulations. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 711–716, 2003.
- [9] Koen Lenaerts and José A Gutiérrez-Fons. To say what the law of the EU is: methods of interpretation and the European Court of Justice. *Colum. J. Eur. L.*, 20:3, 2013.
- [10] Gianmaria Ajani, Guido Boella, Luigi Di Caro, Livio Robaldo, Llio Humphreys, Sabrina Praduroux, Piercarlo Rossi, and Andrea Violato. The European legal taxonomy syllabus: a multi-lingual, multi-level ontology framework to untangle the web of European legal terminology. *Applied Ontology*, 11(4):325–375, 2016.
- [11] Rohan Nanda, Luigi Di Caro, Guido Boella, Hristo Konstantinov, Tenyo Tyankov, Daniel Traykov, Hristo Hristov, Francesco Costamagna, Llio Humphreys, Livio Robaldo, et al. A unifying similarity measure for automated identification of national implementations of European Union directives. In *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*, pages 149–158, 2017.
- [12] Rohan Nanda, Giovanni Siragusa, Luigi Di Caro, Guido Boella, Lorenzo Grossio, Marco Gerbaudo, and Francesco Costamagna. Unsupervised and supervised text similarity systems for automated identification of national implementing measures of European directives. *Artificial Intelligence and Law*, 27(2):199–225, 2019.
- [13] Ali Sadeghian, Lakshman Sundaram, D Wang, W Hamilton, Karl Branting, and Craig Pfeifer. Semantic edge labeling over legal citation graphs. In *Proceedings of the workshop on legal text, document, and corpus analytics (LTDCA-2016)*, pages 70–75, 2016.
- [14] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. In *NeurIPS EMC² Workshop*, 2019.
- [15] Daniela Tiscornia and Fabrizio Turchi. Formalization of legislative documents based on a functional model. In *Proceedings of the 6th international conference on Artificial intelligence and law*, pages 63–71, 1997.
- [16] Bernhard Waltl, Jörg Landthaler, and Florian Matthes. Differentiation and empirical analysis of reference types in legal documents. In *JURIX*, pages 211–214, 2016.
- [17] Aleksandra Čavoški. Interaction of law and language in the EU: Challenges of translating in multilingual environment.” *Perspectives: The journal of specialised translation* 27:58–74. 2017