Summary of the Workshop Legal Information Retrieval meets Artificial Intelligence (LIRAI'23)

Sabine Wehnert[1,2], Manuel Fiorelli[3], Davide Picca[4], Ernesto William De Luca[1,2] and Armando Stellato[3]

[1]*Otto von Guericke University Magdeburg, Germany*
[2]*Leibinz Institute for Educational Media | Georg Eckert Institute, Braunschweig, Germany*
[3]*Tor Vergata University of Rome, Italy*
[4]*University of Lausanne, Switzerland*

Abstract

This paper summarizes the workshop Legal Information Retrieval meets Artificial Intelligence (LIRAI'23), held in Rome, Italy and co-located with the ACM Hypertext conference. In this workshop, 6 publications on the intersection between Legal Information Retrieval and Artificial Intelligence were presented. Enrico Francesconi held a keynote on Profiles of Knowledge Representation and Reasoning for Legal Information Retrieval and Compliance Checking. Overall, the workshop fostered fruitful discussions in various current research areas and challenges, such as the lack of datasets in the legal domain that are dedicated to the evaluation of explainability, the use of Semantic Web together with state-of-the-art transformer architectures, as well as the use of Large Language Models in the legal domain.

Keywords

Legal Informatics, Legal Information Retrieval, Legal Knowledge Representation, Legal Text Mining, Legal Compliance, FAIRness, Semantic Web, Linguistic Legal Linked Open Data, Explainable AI, High-Recall Retrieval

1. Introduction

In the workshop Legal Information Retrieval meets Artificial Intelligence (LIRAI'23), we discussed the complexity of legal systems and the challenges in navigating and understanding legal documents due to their evolving nature and interconnections. We explored the intersection of Legal AI and hypertext, emphasizing the difficulties in retrieving, comprehending, and predicting relationships between legal documents.

Core topics included hypertext-based legal systems, information retrieval and extraction, knowledge graphs, relation extraction, and explainability in legal document retrieval. More

details on the scope of LIRAI can be found in the workshop proposal [1]. This workshop highlighted legal informatics as an interdisciplinary field that leverages information technology to improve legal processes and emphasizes the importance of adopting FAIR practices for publishing legal knowledge.

The relevance of the workshop to hypertext communities[1] lies in the evolving nature of legal knowledge, now being digitized and made available as hypertexts. With the adoption of machine-readable formats and Semantic Web technologies, legal knowledge not only exists on the Web but is an integral part of it, making legal documents a unique case study within hypertexts. The workshop explored these distinct characteristics through the contributions of 6 accepted papers (out of 8 submissions), a keynote and a final discussion. We allowed for different kinds of papers: industrial, demo, position, discussion on emerging topics, short, late-breaking and full papers. The topics of interest and submission criteria for the different paper types can be accessed via the LIRAI'23 website[2].

We extracted several core themes from the accepted paper abstracts, which we visualize with the word cloud in Figure 1. As expected, many keywords are related to information retrieval (e.g., similarity, extraction, keyword), whereas some are closer to general artificial intelligence themes (e.g., models, summarization, adaptation). What we also notice from this birds-eye view are the aspects that have been analyzed (e.g., topic, sentiment), as well as the types of legal documents (e.g., articles, constitutions, patents, court documents).

Figure 1: Word Cloud extracted from accepted paper abstracts visualizing core themes of the contributions.

In the remainder of this summary paper, we describe the keynote speech and the accepted papers of the LIRAI'23 workshop. Then, we recap insights and future directions that were

---

[1]LIRAI was co-located with the ACM Hypertext 2023.
[2]https://sites.google.com/view/lirai-2023/call-for-papers

pointed out during the general discussion, before concluding the event.

2. Contributions

In this section, we present the contributions to the LIRAI'23 workshop. Starting with an abstract of the keynote by Enrico Francesconi, we summarize each accepted paper's contribution.

The keynote speech was held by Enrico Francesconi, who is a Research Director at IGSG-CNR, the Institute for Legal Informatics and Judicial Studies of the National Research Council of Italy and, currently, he is Policy Officer of the European Parliament. In his talk, he focused on creating rules that computers can understand and act upon in the legal field, enabling advanced information services with automatic reasoning abilities. He presented an approach for representing legal knowledge and reasoning using the Semantic Web framework. This method distinguishes between provisions and norms and offers reasoning capabilities like Hohfeldian reasoning for complex legal information retrieval and compliance checking for deontic norms. The approach can handle norm defeasibility and is implemented using specific parts of OWL 2, ensuring computability, while legal reasoning is executed through existing decidable reasoners.

The paper proposed by Dessi et al. primarily aims to enhance the process of extracting relevant keywords from patent documents, a task increasingly important due to the growing volume of patent data. The authors address this challenge by proposing DeepKEA, a novel deep learning model. DeepKEA operates through two interconnected modules. The first module focuses on generating training data, where noun phrases are extracted from the abstracts and claims of patent documents and validated by patent experts. This process ensures that the initial set of terms is closely related to the patent content and inventions described. The second module involves a deep neural network, which is trained using these expert-validated terms and the original patent texts. This network is responsible for the actual keyword extraction from patents. The paper's findings highlight the potential of applying advanced AI techniques, specifically deep learning, in the domain of information retrieval and patent document analysis.

The work by Bauer et al. presents a novel approach to extractive summarization of U.S. Supreme Court opinions. The aim is to address the challenge posed by the typically lengthy nature of these documents, which makes them difficult for both legal professionals and the general public to quickly comprehend. The study employs a dataset of 430K U.S. court opinions with key passages annotated, utilizing this data to train deep learning models for summarization. Among these models, a reinforcement learning-based model named MemSum stands out for its effectiveness, demonstrating superior performance over other approaches, including transformer-based models.

The paper proposed by Simeri and Tagarelli focuses on improving the retrieval of articles from the General Data Protection Regulation (GDPR). GDPR is a critical European regulation that impacts data protection and privacy across the European Union and beyond. The complexity of GDPR's legal texts poses significant challenges for various stakeholders, including government agencies, law firms, legal professionals, and citizens. To address these challenges, the paper proposes an approach using pre-trained language models (PLMs). The authors employ both domain-general and domain-specific pre-trained BERT models, enhancing them through self-supervised task-adaptive pre-training. This approach incorporates stages with or without data

enrichment based on recitals, adding a layer of specificity to the models. The study aims to showcase the effectiveness of PLMs in navigating the intricate legal framework of the GDPR.

Bertillo et al. introduce a technique involving three components (Speaker Diarization, Speech2Text, and Semantic Textual Similarity) designed to enhance a written report about parliamentary debates in Italy with the corresponding video reference. The process involves Speaker Diarization to identify speakers, Speech2Text conversion for transcription, and Semantic Textual Similarity to match and align transcriptions with reports. The resulting Video Table of Contents (VTOC) files allow dynamic navigation of video content, aiding user interaction with the Senate's archive. The evaluation of the system's performance, utilizing precision, recall, and F-measure metrics, demonstrates promising results in identifying semantically similar sentence pairs and generating VTOC files accurately. However, further improvements are suggested for each step, such as enhancing Speech2Text accuracy, refining Speaker Diarization to handle overlapping voices better, and exploring advanced models for Semantic Textual Similarity. Overall, the system shows potential for automating transcription and indexing of Senate sittings.

Greco and Tagarelli share findings on the automated identification of similarities within the constitutions of European countries, specifically employing transformer-based language models. They focus on the important topic of the rights and duties of citizens and use an existing labeling scheme, dividing their data into macro-topics and micro-topics. In their results, we learn that lexical analysis struggles to capture fine-grained similarities between countries, while models that were pre-trained on legal data excel in recognizing similarities but struggle with subtle distinctions. While sentence-transformers show promise in that regard, even they face challenges with closely related micro-topics. The analysis also reveals discrepancies in models' abilities to discern topics within the same macro-category. Models like all-mpnet-base-v1 perform best, displaying clear distinctions among micro-topics. Future enhancements could involve fine-tuning models on constitution data and exploring a broader range of legal models to understand their capabilities better.

Wehnert et al. analyze sentiment and subjectivity in the legal facts of Swiss Federal Supreme Court cases. Their work covers all three languages that the court operates in: German, French, and Italian. They employ out-of-the-box classifiers or dictionaries to detect subjectivity and sentiment. Sharing preliminary results, they make a point about the very limited availability of resources for sentiment analysis, especially in the legal field. They demonstrate several examples in which the available tools make wrong predictions. Furthermore, they show that among the four methods they applied per language, there was almost no consensus, as measured via inter-annotator agreement. Most tools are trained on product reviews or social media text, making them in many cases unsuitable for legal data. The study assumed that there may not be any polarity or subjectivity in the facts that are used for legal decision-making. However, there were several instances in the data that show polarity or subjectivity. Hence, for detecting these cases more reliably, future work on domain-specific classifiers is required.

3. Insights and Future Directions

The workshop provided a stimulating forum for discussing recent advancements in applying NLP and AI technologies in the legal domain. Key insights emerged from the presented papers, along with enriching debates on various aspects of this interdisciplinary field. Here, we want to put forward some general considerations on the main points highlighted during the workshop, including the challenges and opportunities in leveraging NLP for legal applications.

The Challenge of Dataset Availability for Evaluating Explainability   A recurring theme in the workshop was the acute shortage of specialized datasets in the legal domain, particularly those that could aid in evaluating the explainability of AI models. Explainability is crucial in legal contexts, given the need for transparency and accountability in decisions derived from AI systems. The presented papers and discussions, while pioneering in their approaches to legal text processing, also highlighted this gap. The lack of such datasets hinders the development of models that not only perform with high accuracy but also provide understandable decision-making processes. This gap calls for a collaborative effort in the legal and AI research communities to create and share datasets that emphasize not just performance metrics but also the explainability aspects of AI applications in law. For instance, such a dataset can consist of providing highlights of deciding text passages (e.g., in the dataset [2] for a decision rationale extraction task). In some cases providing both, pro and con rationale, can be helpful. An example for this is the work by Yu et al. [3], who publish a dataset with manual annotations by legal experts for both kinds of rationale, along with their alignment and natural language explanations. Nowadays it is often the case that built-in explainability may not be possible due to the nature of the algorithm (e.g., deep learning-based solutions). Instead of following an Explainable AI approach, we may resort to Justifiable AI [4], where we use a retrieval module for extracting supporting and contradicting evidence from a given dataset or trustworthy external sources to fact-check a model's output, so that the users can make an informed decision about the AI's output themselves.

Integrating Semantic Web with State-of-the-Art Transformer Architectures   Another significant point of discussion was the integration of Semantic Web technologies with advanced transformer architectures. This integration promises to enhance the understanding and processing of legal documents. Semantic web technologies facilitate a more nuanced and structured representation of legal knowledge, enabling AI models to better capture the intricacies of legal language and reasoning. When combined with the powerful contextual understanding capabilities of transformer models, as seen in the papers presented, there is a potential to achieve more sophisticated and accurate legal document analysis. This synergy could lead to more efficient legal research, improved case prediction, and enhanced accessibility of legal information. One example for a joint use of those technologies is the work by Kim et al. [5], who compare legal statute entailment classification outputs of BERT and an approach based on a syntactic parser, article segmentation and negation detection called SYN. Consensus was reached if outputs matched; otherwise, semantic codes from the query and relevant article were checked using the Kadokawa Thesaurus Hierarchy. Shaheen et al. [6] use the EuroVoc taxonomy to perform stratified splits and to create reduced label sets in a n large multi-label text classification task.

Exploring the Use of Large Language Models in Legal Applications   A last point that has been treated during the workshop delved into the potential applications of Large Language Models (LLMs) in the legal domain. The versatility of LLMs, as demonstrated by their success in various fields, positions them as valuable tools for legal text analysis and generation. The papers showcased the effectiveness of LLMs in tasks like summarization and information retrieval within legal documents. However, the discussions also brought forth concerns regarding the biases and the need for domain-specific tuning of LLMs for legal applications. Given the sensitive nature of legal text and the necessity for high accuracy and reliability, the application of LLMs in law requires careful consideration and tailored approaches. As discussed before, the outputs of an LLM may need to be at least justifiable by incorporating a fact-checking mechanism [4], because interpretability and transparency are key in many legal applications. Further pointers on interpretability and knowledge injection in transformers and transformer-based architectures such as LLMs can be found in the systematic study by Greco and Tagarelli [7]

The workshop underscored the significant strides being made in applying NLP to the legal domain, while also highlighting critical areas for future research and development. The integration of AI with legal information retrieval is not without its challenges, but the potential benefits it offers in terms of accessibility, efficiency, and enhanced understanding of legal texts are huge. As the field progresses, it will be crucial to address the current limitations, particularly around the already mentioned critical points such as dataset availability, the integration of semantic technologies, and the use of LLMs, to fully harness the power of AI in legal contexts.

4. Conclusion

In conclusion, the LIRAI'23 workshop at the ACM Hypertext conference provided valuable insights into the intersection of Legal Information Retrieval and Artificial Intelligence. The six accepted papers and keynote by Enrico Francesconi addressed diverse challenges in legal document processing. Key themes during the discussions included the scarcity of datasets for evaluating AI explainability in the legal domain, the integration of Semantic Web technologies with transformer architectures, and the promising yet cautiously approached use of Large Language Models (LLMs) in legal applications. The workshop highlighted the need for collaborative efforts in dataset creation, the potential of semantic technology integration, and the importance of tailored approaches for LLMs in legal contexts. Overall, LIRAI'23 underscored the dynamic evolution of Information Retrieval and AI in the legal domain, while identifying critical areas for future research and development.

Acknowledgments

Members Of The Program Committee:

- Tommaso Agnoloni, ITTIG-CNR, Italy
- Ilaria Angela Amantea, University of Turin, Italy
- Vito Walter Anelli, Politecnico di Bari, Italy
- Dennis Aumiller, Heidelberg University, Germany
- Valerio Basile, University of Turin, Italy
- Luigi Di Caro, University of Turin, Italy
- Harshvardhan J. Pandit, Dublin City University, Ireland
- Mi-Young Kim, University of Alberta, Canada
- Rūta Liepiņa, University of Bologna, Italy
- Carlo Marchetti, Senate of the Republic, Italy
- Patricia Martín-Chozas, Universidad Politécnica de Madrid, Spain
- Elena Montiel-Ponsoda, Universidad Politécnica de Madrid, Spain
- Jack Mumford, University of Liverpool, UK
- Monica Palmirani, University of Bologna, Italy
- Ginevra Peruginelli, ITTIG-CNR, Italy
- Ken Satoh, National Institute of Informatics and Sokendai, Japan
- Emilio Sulis, University of Turin, Italy
- Andrea Tagarelli, University of Calabria, Italy
- Marc van Opijnen, Publications Office of the Netherlands, The Netherlands
- Eugene Yang, Human Language Technology Center of Excellence, Johns Hopkins University, USA
- Masaharu Yoshioka, Hokkaido University, Japan
- Tomasz Zurek, T.M.C. Asser Institute, University of Amsterdam

References

[1] E. W. D. Luca, M. Fiorelli, D. Picca, A. Stellato, S. Wehnert, Legal information retrieval meets artificial intelligence (LIRAI), in: Proceedings of the 34th ACM Conference on Hypertext and Social Media, HT 2023, Rome, Italy, September 4-8, 2023, ACM, 2023, pp. 46:1–46:4. doi:10.1145/3603163.3610575.

[2] I. Chalkidis, M. Fergadiotis, D. Tsarapatsanis, N. Aletras, I. Androutsopoulos, P. Malakasiotis, Paragraph-level rationale extraction through regularization: A case study on european court of human rights cases, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tür, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, Association for Computational Linguistics, 2021, pp. 226–241. doi:10.18653/V1/2021.NAACL-MAIN.22.

[3] W. Yu, Z. Sun, J. Xu, Z. Dong, X. Chen, H. Xu, J. Wen, Explainable legal case matching via inverse optimal transport-based rationale extraction, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 657–668. doi:10.1145/3477495.3531974.

[4] S. Wehnert, Justifiable artificial intelligence: Engineering large language models for legal applications, CoRR abs/2311.15716 (2023). doi:10.48550/ARXIV.2311.15716. arXiv:2311.15716.

[5] M. Kim, J. Rabelo, K. Okeke, R. Goebel, Legal information retrieval and entailment based on bm25, transformer and semantic thesaurus methods, Rev. Socionetwork Strateg. 16 (2022) 157–174. doi:10.1007/S12626-022-00103-1.

[6] Z. Shaheen, G. Wohlgenannt, E. Filtz, Large scale legal text classification using transformer models, CoRR abs/2010.12871 (2020). arXiv:2010.12871.

[7] C. M. Greco, A. Tagarelli, Bringing order into the realm of transformer-based language models for artificial intelligence and law, CoRR abs/2308.05502 (2023). doi:10.48550/ARXIV.2308.05502. arXiv:2308.05502.