ESELS

Welcome to the ESELS Newsletter

As of 2024, the **European Society for Empirical Legal Studies** publishes its newsletter twice a year. The aim is to inform ESELS members about the Society's activities and conferences as well as highlight the work of ESELS members.

In this second edition, we take a look back at our very successful **ESELS 2024 Annual Conference in Elche**.
We also look forward to our upcoming Annual Conferences at the **University of Toulouse (France)** in 2025 and the **University of Ljubljana (Slovenia)** in 2026. (Keep an eye out for our **call for abstracts** in early 2025!)
We are also happy to share the **results of the 2024 elections** that took place during the Annual Conference in Elche.
Moreover, the Board has launched a **call for proposals** by ESELS members for **ESELS working groups**, more on that below.
Finally, **Gaspar Dugac and Tilmann Altwicker** (University of Zürich) shed light on the phenomenon of Large Language Models and potential role for empirical legal studies.

Happy reading!

Do you have ideas, suggestions or questions about this newsletter? Please contact the ESELS Newsletter editor Kyra Wigard

Looking back: ESELS 2024 Annual Conference at Universidad Miguel Hernández, Elche

On **20-21 June 2024** the highly successful Annual Conference of ESELS took place in the city of **Elche** (Spain), hosted by the **Universidad Miguel Hernández**. For two days more than 120 scholars working in the field of Empirical Legal Studies from all over Europe and around the world convened and contributed to fruitful intellectual exchanges and collaboration on empirical-legal research and teaching.

With illuminating **keynote** speeches by **Juan S. Mora-Sanguinetti**, **Barbora Holá** and **Mathias Siems**, conference participants also attended various interactive **panel sessions** on Empirical Legal Research over two days, in which scholars from different legal backgrounds and disciplines showcased their researching findings and engaged in productive discussions with their ELS colleagues. The city of Elche provided a beautiful backdrop for our discussions and our social gathering with drinks at the pool of the beautiful, palm-surrounded **Hotel Huerto del Cura**, followed by a wonderful dinner featuring many delicious Spanish specialties.

The closing ceremony was chaired by past President **Catrien Bijleveld**, delivering thankful words to the outstanding local organization for making this great event possible. On behalf of the local organization **Fernando Miró Llinares** handed over the ESELS flag to Executive Secretary Erik Wesselius, to be displayed again during the next **Annual Conference in Toulouse.**

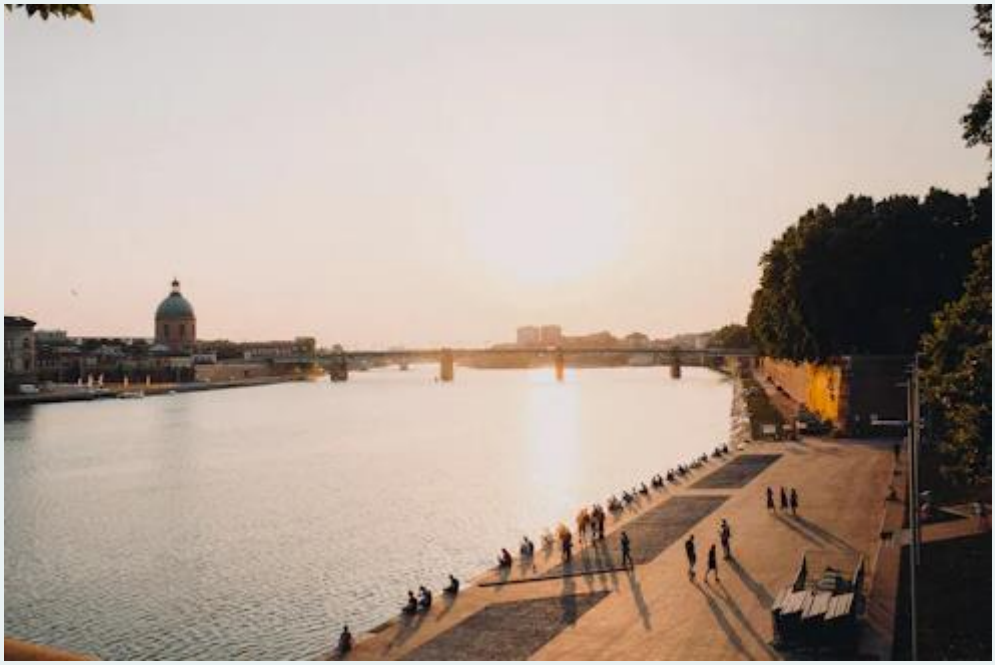
Read the conference report

Looking ahead:
ESELS 2025 Annual Conference

The next Annual ESELS Conference will take place on 19-20 June 2025 and will be hosted in the heart of Europe by the **Université Toulouse Capitole** – one of the oldest universities in Europe (founded in 1229).

The Université Toulouse Capitole is known for its commitment to Empirical Legal Research, having organized the 2023 International Conference on ELS and 2023 and 2024 workshops on Empirical Legal Research.

We look forward to seeing many of you there!



Newly-elected board members and ESELS President

During the Annual Conference, the ESELS General Assembly convened and attending ESELS members elected **Tilmann Altwicker** as ESELS Future President (2025/26) and **Rita Gsenger** as ESELS Board Member.

Now past President **Catrien Bijleveld** handed over the reigns of the society to **Urška Šadl** who will be our President in 2024/25.

Do you want to know more about our 2024/25 Board? Please take a look below!

Get to know the 2024/25 ESELS Board

ESELS Working Groups

One of the results of the 2024 General Assembly is the launch of a call for proposals for **ESELS working groups**. Our Society welcomes the creation of **thematic** working groups in Empirical Legal Studies.

Working groups are expected to contribute to fruitful exchange and stimulation of empirical-legal research in areas relevant to a European audience of scholars.
Members of working groups must be members of our Society and are expected to
1) hold regular meetings or exchanges and
2) to organize at least one pre-arranged panel at each year's conference.

Some examples of topics potentially interesting for ESELS working groups are: "Empirical Comparative Law", "Large Language Models in Law", "Empirical Constitutional Studies" or "AI and the Law"…

Proposals for ESELS working groups will be assessed and decided upon by the ESELS Board.

If you have any ideas or proposal, please submit them through the link below.

More information and proposal submissions

EJELS

EUROPEAN JOURNAL FOR EMPIRICAL LEGAL STUDIES

More information and article submissions

Boosting Empirical Legal Studies using Large Language Models

by Gaspar Dugac and Tilmann Altwicker

August 2024

1 The Environment of LLMs and Law

It is without a doubt that the field of Empirical Legal Studies (ELS) is currently growing and diversifying. While the adoption of quantitative methods has considerably propelled the empirical study of the law, there still seems to be some hesitation to adopt Large Language Models (LLMs) to assist empirical legal research. With this brief post, we wish to explain the added value of integrating LLMs into the methodological toolkit of empirical legal scholars.

Given that legal studies inherently deal with analyzing text, it comes as no surprise that some researchers have been operating at the intersection of law and natural language processing for a while (Chalkidis et al., 2020; Merchant and Pande, 2018; Nguyen et al., 2018). However, with the advent of LLMs, we observed an increased interest in how these tools affect the legal field, both from academia, as well as from popular culture. Media widely reported the success of ChatGPT in passing the bar exam and there have been many papers evaluating LLMs on similar tasks (Choi et al., 2023; Katz et al., 2023). Since then, many scholars have been trying to discover where the models' limitations lie when it comes to solving legal tasks. Trozze et al. (2023) studied the models' reasoning and legal drafting capabilities on security cases involving cryptocurrencies, Nayet al. (2023) investigated the reasoning capabilities of several models in applying tax law, and Savelka (2023) evaluated the models' performance in performing zero-shot semantic annotation on text excerpts from a variety of legal documents. The overwhelming number of papers evaluating models on various legal tasks has prompted the consolidation of these tasks into a benchmark, LegalBench, which can be used to evaluate any model on a manually-curated selection of legal tasks (Guha et al., 2023).

2 Over promises and the realistic position of LLMs within ELS

LLMs have become a focal point of many scientific fields, not exclusively legal studies. It would not be untrue to state that the field has been saturated with the promises of what these models can and cannot do. Despite their potential, the use of LLMs in ELS is not without challenges. Research and practice have both shown the limitations of these models, including but not limited to, hallucinations (Ji et al., 2023; Ye and Durrett, 2022), interpretability (Benderet al., 2021; Shen et al., 2023; Saha et al., 2023), and security (Yao et al., 2024). Hence, caution in interpreting results generated by these models is necessary in legal contexts where transparency and explainability are crucial.

Our view is that the current position of these models is to complement classical quantitative methods and reducing manual labour, especially in the early stages of an ELS research project. As empirical researchers, we rely on data which usually difficult to obtain. Data collection and annotation require financial resources, but more importantly they take time. Oftentimes, the research question we wish to investigate does not have existing data. Training human annotators and labeling vast amounts of data is a time-intensive task. One could use basic machine learning methods to automate the extraction of simple variables of interest such as names, dates, etc. However, there is a much larger collection of variables which are harder to identify, and which might appear implicitly in the text. This is precisely the setting where LLMs could be of assistance- as annotators (see above-mentioned Savelka (2023)). In this way, the use of LLMs could boost ELS.

We have observed, from our own research, that LLMs can be used to annotate not only simple variables, but also hard, implicit variables from text. For example, in the past, we evaluated classical machine learning methods on classifying the use of legal interpretation techniques (e.g. teleological interpretation) present in text, say, a judgment. However, this process required annotating thousands of text excerpts to be used as training data. The results were quite impressive, but the entire process proved itself to be too resource-intensive to be applied in the future. Consequently, we pivoted to methods which require minimal training data such as LLMs. Our research, presented at the 2024 ESELS Conference in Elche, showed several LLMs have decent, albeit still limited, capabilities in annotating legal interpretations present in text, a hard variable to automatically annotate, with minimal training data (zero- and 3-shot). Since then, better models have been released and their performance on this task has only improved.

Classical statistics is not going to disappear, though. As scientists we wish to estimate certain effects or associations, and we wish to have some statistical guarantees for the results we obtain. This is the toolbox of statistical inference and remains the main way we conduct empirical legal research. However, this approach is limited to the data we have available; and for a long time, data availability will remain an issue in ELS. Oftentimes, in ELS we relegate investigation to vignette studies, or similar, simply due to the lack of data. LLMs

might be a tool researchers can use to bridge the gap between what we want to research and what we are able to research based on current data constraints.

3 How should ELS researchers proceed?

We suggest ELS researchers to approach LLMs with a combination of curiosity, critical thinking, and a collaborative mindset. We recognize the barriers to entry when it comes to LLMs, both from a theoretical point-of-view, as well as an applied one. However, there has never been as abundant a supply of learning resources as there is today. Keeping up with these developments is an investment into the future of ELS and will be a necessity to progress as researchers. Beyond familiarising themselves with the necessary programming languages such as Python and understanding the underlying structure of LLMs on a broad level, ELS researchers should experiment with LLMs on small-scale projects. Hands-on experience is essential for learning how to effectively use LLMs. These projects will provide practical insights into the capabilities and limitations of LLMs. Additionally, they offer a low-risk environment to experiment with different models, parameters, and datasets. As researchers gain confidence, they can gradually incorporate LLMs into larger ELS projects.

Once researchers gain basic experience in using LLMs, they can employ further methods to obtain more reliable and consistent results, such as self consistency and better prompting strategies (Wang et al., 2023; Kojima et al., 2023). Some researchers have already employed prompting strategies using their knowledge of the legal domain. Jiang and Yang (2023) used legal syllogism to construct chain-of-thought prompts, while Yu et al. (2022) derived prompts from specific legal reasoning techniques such as IRAC (Issue, Rule, Application, Conclusion). Future research could focus on developing domain-specific models tailored to legal research, for example using fine-tuning (Hu et al., 2021; Dettmers et al., 2023). Moreover, one could use Retrieval-Augmented Generation (RAG) (Gao et al., 2024) to connect LLMs to a legal studies oriented knowledge base in order to improve the accuracy and credibility of model-generated outputs.

Most importantly, given the fast-paced development of these tools, combined with the relatively recent development of European ELS, we stress the critical importance of collaboration and community engagement for staying informed about the latest developments and their applications in legal research. This could be through conferences, workshops, seminars, or working groups, and should include people with various backgrounds. Engaging with the broader legal and data science communities allows researchers to learn from others' experiences, discuss their own findings, and stay updated on cutting-edge research approaches. Additionally, collaboration with data scientists and computer scientists can provide valuable technical expertise that complements legal knowledge.

In conclusion, LLMs offer significant promise for ELS by enhancing the efficiency, accuracy, and scope of empirical legal research. While challenges remain, the careful and responsible application of these models has the potential to transform the way legal scholars and practitioners engage with legal data.

References

Bender, E. M., T. Gebru, A. McMillan-Major, and S. Shmitchell (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, New York, NY, USA, pp. 610–623. Association for Computing Machinery.

Chalkidis, I., M. Fergadiotis, P. Malakasiotis, N. Aletras, and I. Androutsopoulos (2020, November). LEGAL-BERT: The muppets straight out of law school. In T. Cohn, Y. He, and Y. Liu (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2020*, Online, pp. 2898–2904. Association for Computational Linguistics.

Choi, J. H., K. E. Hickman, A. Monahan, and D. B. Schwarcz (2023). Chatgpt goes to law school. *71 Journal of Legal Education 387*.

Dettmers, T., A. Pagnoni, A. Holtzman, and L. Zettlemoyer (2023). Qlora: Efficient finetuning of quantized llms.

Gao, Y., Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, M. Wang, and H. Wang (2024). Retrieval-augmented generation for large language models: A survey.

Guha, N., J. Nyarko, D. E. Ho, C. R´e, A. Chilton, A. Narayana, A. Chohlas-Wood, A. Peters, B. Waldon, D. N. Rockmore, D. Zambrano, D. Talisman, E.Hoque, F. Surani, F. Fagan, G. Sarfaty, G. M. Dickinson, H. Porat, J. Hegland, J.Wu, J. Nudell, J. Niklaus, J. Nay, J. H. Choi, K. Tobia, M. Hagan, M.Ma, M. Livermore, N. Rasumov-Rahe, N. Holzenberger, N. Kolt, P. Henderson, S. Rehaag, S. Goel, S. Gao, S. Williams, S. Gandhi, T. Zur, V. Iyer, andZ. Li (2023). Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models.

Hu, E. J., Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W.Chen (2021). Lora: Low-rank adaptation of large language models.

Ji, Z., N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung (2023, mar). Survey of hallucination in natural language generation. *ACM Computing Surveys 55* (12), 1–38.

Jiang, C. and X. Yang (2023). Legal syllogism prompting: Teaching large language models for legal judgment prediction.

Katz, D. M., M. J. Bommarito, S. Gao, and P. Arredondo (2023). Gpt-4 passes thebar exam. *SSRN Electronic Journal*.

Kojima, T., S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa (2023). Large language models are zero-shot reasoners.

Merchant, K. and Y. Pande (2018). Nlp based latent semantic analysis for legal text summarization. In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1803–1807.

Nay, J. J., D. Karamardian, S. B. Lawsky, W. Tao, M. Bhat, R. Jain, A. T. Lee, J.H. Choi, and J. Kasai (2023). Large language models as tax attorneys: A case study in legal capabilities emergence.

Nguyen, T.-S., L.-M. Nguyen, S. Tojo, K. Satoh, and A. Shimazu (2018, jun). Recurrent neural network-based models for recognizing requisite and effectuation parts in legal texts. *Artif. Intell. Law 26* (2),169–199.

Saha, T., D. Ganguly, S. Saha, and P. Mitra (2023). Workshop on large language models' interpretability and trustworthiness (llmit). In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, CIKM '23, New York, NY, USA, pp. 5290–5293. Association for Computing Machinery.

Savelka, J. (2023, jun). Unlocking practical applications in legal domain: Evaluation of gpt for zero-shot semantic annotation of legal texts. In *Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law*, ICAIL2023. ACM.

Shen, X., Z. Chen, M. Backes, and Y. Zhang (2023). In chatgpt we trust? measuring and characterizing the reliability of chatgpt.

Trozze, A., T. Davies, and B. Kleinberg (2023). Large language models in crypto currency securities cases: Can chatgpt replace lawyers?

Wang, X., J. Wei, D. Schuurmans, Q. Le, E. Chi, S. Narang, A. Chowdhery, and D. Zhou (2023). Self-consistency improves chain of thought reasoning in language models.

Yao, Y., J. Duan, K. Xu, Y. Cai, Z. Sun, and Y. Zhang (2024). A survey on large language model (llm) security and privacy: The good, the bad, and the ugly.

Ye, X. and G. Durrett (2022). The unreliability of explanations in few-shot prompting for textual reasoning. In A. H. Oh, A. Agarwal, D. Belgrave, and K.Cho (Eds.), *Advances in Neural Information Processing Systems*.

Yu, F., L. Quartey, and F. Schilder (2022). Legal prompting: Teaching a language model to think like a lawyer.