Textual Insights: What Can Computers Teach Legal Scholars About Law?

Johan Lindholm

ABSTRACT

Legal research has historically relied on the manual and systematic study of authoritative texts, a methodology that has remained largely unchanged despite technological advancements. However, recent developments in natural language processing and other data-driven approaches present new opportunities for legal scholars. This essay examines whether and how these computational tools can complement doctrinal approaches and explores the potential of computational methods to enhance and transform legal scholarship. In emphasizing the compatibility of computational and doctrinal approaches, it argues that by integrating these approaches, legal scholars can make scientific discoveries beyond the scope of either method alone. The essay concludes by outlining the steps necessary for legal scholarship to fully embrace and benefit from these emerging technologies.

Keywords: legal scholarship, computational methods, empirical legal studies, natural language processing, large language models.

1. INTRODUCTION

For centuries, lawyers, judges, law students, and legal scholars alike would gather in law libraries to 'do legal research'. By 'doing legal research' they meant roughly the same thing: to carefully study authoritative texts in order to determine what the law governing a particular issue 'is', that is to say what is commonly referred to as conducting doctrinal legal research. They all used more or less the same methodology which centered around the manual and systematic reading of authoritative legal sources, although they did this for slightly different motives, e.g. to advise and represent clients, to justify judgments, to learn the law, and to improve the legal system.[1]

At the turn of the twentieth century, legal realists on both sides of the Atlantic challenged traditional conceptions of law. Rejecting what they viewed as the metaphysical elements of law, they instead saw it as an inherently social phenomenon. For instance, under this perspective, a statement about what the law is could be understood as a prediction of how judges will apply it to specific facts.[2] One could imagine that this would change how

legal research was conducted. In America, this was also to some extent the case. Realists accepting law as inherently entangled in a messy social and political reality[3] would pave the way for a methodological turn towards the empirical.[4] This was less true in Europe where legal scholars and practising lawyers largely continued going to the law library to find and read the authoritative legal documents, largely in the same way as each other, and largely using the same methods as before.

Even the technical revolution that took place in the 1990s did not fundamentally change how legal research was conducted. The main contribution that general access to affordable personal computers, the invention of CD-roms, and even the introduction of the Internet made to legal research was that the authoritative sources traditionally studied now could be accessed digitally while the paper versions collected dust on library shelves.[5] This

---

1. See e.g. Rob van Gestel and Hans-Wolfgang Micklitz, 'Why Methods Matter in European Legal Scholarship: Methods in European Legal Scholarship' (2014) 20 European Law Journal 292; Terry Hutchinson and Nigel Duncan, 'Defining and Describing What We Do: Doctrinal Legal Research' (2012) 17 Deakin Law Review 83.

2. Alf Ross, *On Law and Justice* (Jakob vH Holtermann ed, Uta Bindreiter tr, Oxford University Press 2019) 156; see also Éric Millard, 'Alf Ross and Realist Conceptions of Legislation' in Pierre Brunet, Éric Millard and Patricia Mindus (eds), *The Theory and Practice of Legislation* (Hart

Publishing 2013); Oliver Wendell Holmes, 'The Path of the Law' (1997) 110 Harvard Law Review 991, 994 ("The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law.").

3. Gregory S Alexander, 'Comparing the Two Legal Realisms-American and Scandinavian' (2002) 50 The American Journal of Comparative Law 131, 133.

4. Michael Heise, 'The Past, Present, and Futute of Empirical Legal Scholarship: Judicial Decision Making and the New Empiricism' (2002) 2002 University of Illinois Law Review 819, 822–824.

5. Despite some valiant efforts towards change. For a historical overview, including of the efforts that were made, see Peter Wahlgren, 'The Quest for Scientific Methods: Sociology of Law, Jurimetrics and Legal Informatics' in Håkan Hydén and others (eds), *Combining the Legal and the*

was true for practising lawyers as well as for most legal scholars.

We find ourselves more recently in the middle of another technological revolution. Increasingly easy access to large, accurate, and accessible datasets on law and legal institutions combined with a methodological development that can best be described as dizzying presents legal scholars with a rich toolbox of exciting computational methods at their disposal.[6] The term 'computational methods', as it is used in this contribution, refers to a broad range of data-driven approaches developed in the field of computer science.[7] The development in natural language processing (NLP) methods, in particular the introduction of word embeddings[8] and transformers,[9] deserve special attention as they are, on their face, ideally suited for a text-focused discipline like law.

Will this development change how legal scholarship is done? I will address whether and to what extent legal scholarship[10] can benefit from using computational methods, i.e. what is sometimes referred to as data science in law,[11] law-as-data,[12] or computational legal studies (CLA).[13] The vein of scholarship that I will focus on shares the theoretical and epistemological foundations of many types of empirical legal scholarship,[14] but falls closer to legal informatics with regards to method.[15] My main point is that computational and doctrinal approaches are compatible and that combining them allows for scholarly discoveries beyond the reach of either by themselves.[16] After presenting the reasons for this position, I will discuss the steps needed to move forward.

2. TRACING THE ROOTS OF MISDIRECTED EMPIRICAL SCEPTICISM

It seems that there is considerable scepticism among European legal academics regarding the use of empirical methods in legal research and that this is the source of some tension between scholars that employ empirical methods and those who do not.[17] An example of such scepticism can be found in Hesselink's claim that "[i]f one wants to know what the right answer is to a question of law then empirical research of whatever kind will simply not be helpful".[18] Even more bluntly formulated, Kochenov believes that "while there is law and there is empirical research, doing the latter in order to pretend to say anything about the former... is both methodologically and theoretically dubious, if not nonsensical".[19] While I wholeheartedly believe (i) that it is important to answer the type of research questions that legal scholars have traditionally asked, (ii) that legal scholars by answering such questions fill an important role in society, and (iii) that a thorough understanding of law of the kind that one attains through legal education and training is required in answering those questions,[20] I respectfully disagree

Social in Sociology of Law: An Homage to Reza Banakar (Hart Publishing 2023).

6 Elena Kantorowicz-Reznichenko, 'Computational Methods for Legal Analysis: The Way Forward?' (2021) 14 Erasmus Law Review.

7 Cf. ibid. The terminology in the field is both vast and complicated and, in order to not unnecessarily confuse the reader, I will try to keep it as simple as possible. The type of methods discussed in this contribution includes what is sometimes refereed to as data science and artificial intelligence (AI) methods. This broad category includes, among other things, machine learning (ML) – which *inter alia* includes deep learning – and natural language processing (NLP) – which in turn includes, among other things, large language models (LLMs). It also includes "non-AI" methods, including some of the methods used in quantitative text analysis (QTA). See Bao Kham Chau and Michael A Livermore, 'Studying Judicial Behavior with Text Analysis' in Lee Epstein and others (eds), *Oxford Handbook of Comparative Judicial Behavior* (Online version, Oxford University Press 2024), including so-called text and data mining (TDM), as well as for example network analysis (NA). It does not however include more traditional frequentist statistical methods (not that there is anything wrong with these methods, I use them myself all the time).

8 Tomas Mikolov and others, 'Efficient Estimation of Word Representations in Vector Space' <http://arxiv.org/abs/1301.3781> accessed 29 September 2023. Word embeddings are representations of words in a continuous vector space, where words with similar meanings have similar vector representations.

9 Ashish Vaswani and others, 'Attention Is All You Need' <http://arxiv.org/abs/1706.03762> accessed 17 March 2023. Transformers are context-aware embeddings, i.e. the embedding of a word depends on the context it is used, and is serves as the basis for many state-of-the-art models like BERT and GPT.

10 As discussed in Section 2, I here chose to define 'legal scholarship' broadly based on the knowledge it tries to produce rather than by the methods it (traditionally) uses. I here deliberately do not use the more established term 'legal research', even though they could be synonymous, in order to avoid confusion with the type of legal research that non-scholar lawyers engage in. Although scholarly and non-scholarly legal research *may* significantly overlap with regard to aim, theory, and method, a crucial point of departure for this essay is that they do not *necessarily* do so. In a Swedish context, it would be natural to use the term 'legal science' (*rättsvetenskap*), but I fear that it might spark connotations to and questions about whether legal scholarship is sufficiently scientific, which is not this contributions' subject and might detract from its actual one. Finally, it is also worth clarifying that I do not even entertain the idea that legal scholarship should only use computational methods, nor herein seek to address the appropriateness of using "AI" in law outside the scientific domain, for example automated decision-making.

11 Jinzhe Tan and others, 'Data Science Applications and Implications in Legal Studies: A Perspective Through Topic Modelling' [2023] Journal of Data Science 57, 2.

12 Michael A Livermore and Daniel N Rockmore (eds), *Law as Data: Computation, Text, and the Future of Legal Analysis* (Santa Fe Institute Press 2019); Bao Kham Chau and Michael A Livermore, 'Computational Legal Studies Comes of Age' (2024) 1 European Journal of Empirical Legal Studies; Jens Frankenreiter and Michael A Livermore, 'Computational Methods in Legal Analysis' (2020) 16 Annual Review of Law and Social Science 39, 4–6.

13 Kantorowicz-Reznichenko (n 6).

14 I here consciously refrain from making distinctions between subfields, such as socio-legal studies and law and economics, qualitative and quantitative ELS etc.

15 Thomas Margoni, 'Computational Legal Methods: Text and Data Mining in Intellectual Property Research' in Irene Calboli and Maria Lillà Montagnani (eds), *Handbook of Intellectual Property Research* (Oxford University Press 2021) 490–493; see also Wahlgren (n 5).

16 Cf. Margoni (n 14) 493.

17 Gestel and Micklitz (n 1) 293–297, 300.

18 Martijn Hesselink, 'A European Legal Method? On European Private Law and Scientific Method' (2009) 15 European Law Journal 20, 28.

19 Dimitry Kochenov, 'Counting Swines at a Satan's Ball: Book Review of Jan Zglinski's Europe's Passive Virtues' <http://dx.doi.org/10.2139/ssrn.4086668>.

20 Cf. Richard A Posner, 'The State of Legal Scholarship Today: A Comment on Schlag' (2008) 97 Georgetown Law Journal 845, 854.

with Hesselink's and Kochenov's blanket rejections of the usefulness of empirical methods when it comes to saying something novel about law.[21] Computational methods not only can but have already improved legal scholarship.

It seems that the under-appreciation of computational methods in law can be traced back to certain incorrect ideas and assumptions. There is a presence in legal academia of a certain understanding of what constitutes legal scholarship and that in my opinion is unduly restrictive and scientifically counterproductive. At the root of much traditionalist rejection of empiricism lies a dichotomous distinction between doctrinal legal scholarship that seeks to answer normative questions about the law from a legal-internal perspective and empirical legal scholars that are interested in answering descriptive questions related to law's external effects and relations.[22] This is reflective of a view that it is possible and important to uphold a distinction between doctrinal legal scholars and other scholars interested in law. For example, it is commonplace in legal literature to distinguish between, on the one hand, doctrinal legal research and doctrinalists and, on the other hand, empirical legal studies, empirical social science, and multidisciplinarians.[23]

I think this dichotomous thinking is based on an incorrect belief that legal scholars are primarily interested in normative doctrinalism and deductive analysis, whereas legal scholars in fact frequently make empirical claims.[24] One could even make the case that much (supposedly doctrinal) legal research employs a type of empirical approach in so far that makes a prognosis about how the law will be applied[25] on the basis of what has been said and done in the past.[26] A common strategy employed by legal scholars that make empirical claims – for example about shifts in the law, in legal reasoning, legal culture, or legal institutions – is to provide a few examples. This can essentially be characterized as small-n empirical studies.[27] I have no wish to debate the appropriate terminology for different methodological approaches. My point is that empiricism is not fundamentally alien in legal scholarship and that we should therefore discuss *when*, not *if*, we should use empirical approaches in legal scholarship.

Another assumption concerns what computational and other empirical approaches are and what they can be used for. There is a risk that these ideas are based on an outdated understanding both of what empirical legal scholars do and of the methodological state of the art. It seems that this view stems from the idea that empirical approaches are exclusively capable of saying something about the context surrounding law (the external perspective), and not about the law as such (the internal perspective). I will not deny that much empirical legal scholarship, possibly even the majority, focuses on questions, factors, and phenomena that can be characterised as external to the law. To the extent that there is a dominant view of what empirical legal scholarship can be used for, it might in this way be based on empirical observations of the type of empirical scholarship that has been conducted.

The inclusion of external factors has been promoted as one of the strengths of empirical approaches. By studying, *inter alia*, how people experience interacting with the legal system, how law affects behavior on individual and group levels, the efficacy of policy implemented through law under various conditions, and the micro- and macro-economic impact of legal rules and procedures, empirical approaches to law have produced important knowledge that could not have been attained using exclusively a doctrinal approach.[28] Some of these studies can be characterised as interesting in something different than what has traditionally interested legal scholars. For example, if a law-and-economics scholar using an empirical approach argues in favor of a particular regulatory solution based on market efficiency, this can be seen as distinctly different from doctrinal legal scholarship, something "outside the realm of legal analysis".[29]

That some empirical legal scholars *have historically been* interested in legal-external questions or that some empirical approaches are *unsuitable* for answering legal-internal questions is however irrelevant when it comes to determining whether current state-of-the-art computational approaches are *capable* of answering legal-internal questions.[30] While empirical approaches can be used to uncover information about the context in which law is situated, it does not conversely follow that it is only good for this. I now will provide some concrete and illustrative examples to the contrary.

---

21 In all fairness it should be pointed out that much development has taken place in the decade-and-a-half that has passed since Hesselink made his claim and I do not know to what extent he would stand by it today.

22 See e.g. Hesselink (n 17) 28–39; Gareth Davies, 'The Relationship Between Empirical Legal Studies and Doctrinal Legal Research' (2020) 13 Erasmus Law Review 3, 9.

23 See e.g. Davies (n 20); Sanne Taekema, 'Methodologies of Rule of Law Research: Why Legal Philosophy Needs Empirical and Doctrinal Scholarship' (2021) 40 Law and Philosophy 33; Gestel and Micklitz (n 1).

24 Lee Epstein and Gary King, 'The Rules of Inference' (2002) 69 University of Chicago Law Review 1, 2–4; Gestel and Micklitz (n 1) 302–303.
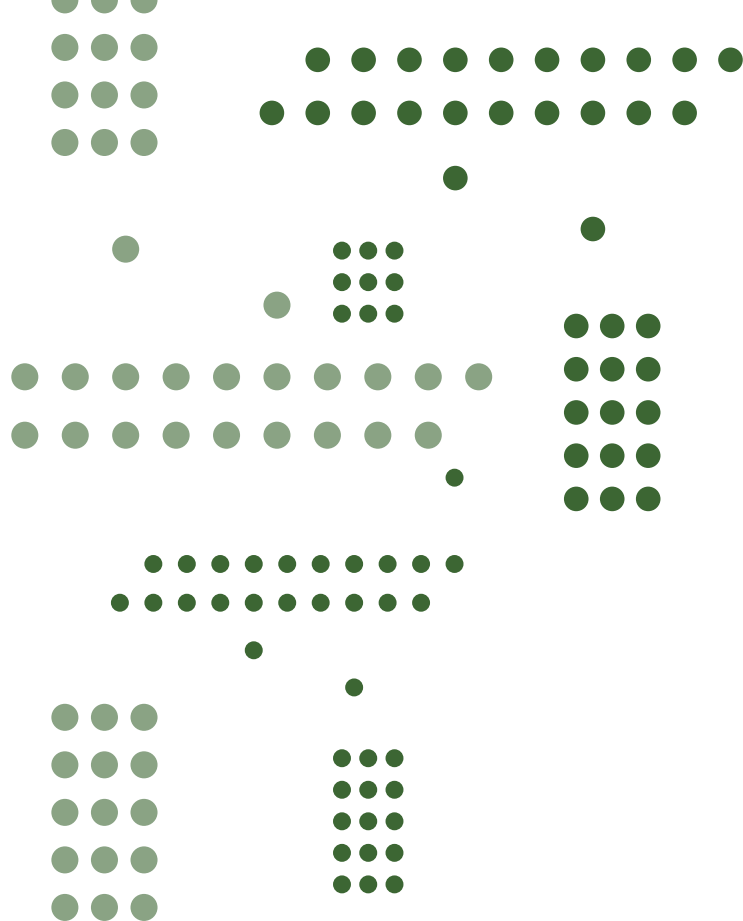
25 See fn 2 and accompanying text.

26 I make this point knowing that not all would share this description.

27 In social science it is commonplace to distinguish between qualitative and quantitative empirical studies. The differences between these approaches and the basis for the distinction, in terms of methodology, epistemology, what they can say about 'reality' etc., are not settled questions. However, one basic difference lies in the number of observations (n) that they study where quantitative studies can be characterized as large-n studies.

28 Deborah R Hensler and Matthew A Gasperetti, 'The Role of Empirical Legal Studies in Legal Scholarship, Legal Education and Policy Making: A US Perspective' in Rob van Gestel, Hans-W Micklitz and Edward L Rubin (eds), *Rethinking Legal Scholarship: A Transatlantic Dialogue* (Cambridge University Press 2017) 469–474.

29 Hesselink (n 17) 30.

30 Much like legal scholarship should not be defined by its dominant methodology, empirical legal studies and the use of empirical methods in legal scholarship should not on principle be limited to legal-external questions.

to hallucinations,[34] and are not perfect at conducting statutory reasoning.[35] While LLMs make mistakes, so do LL.M.s, and it is equally clear that they are not incapable of answering legal questions or nonsensical.[36] I would think that the existence of these LLM-based chatbots should help convince sceptics that computational methods can be useful in legal scholarship, even scholarship that seeks to answer legal-internal questions.

Impressive as they are, LLM-based chatbots are not prone to conduct legal scholarship in the sense that they can generate novel insights about the law. This is clearly the case with the current state of the technology, but it appears to be an inherent limitation of how they are trained. By virtue of being limited to the data that they have been trained on, LLMs are capable of generating information based on what has already been concluded, the type of legal answers that one can find in textbooks. Such answers are clearly not worthless, and because of their ability to generate such information LLMs are valuable tools for scholars conducting research, but they cannot as such produce boundary-pushing research. I shall now provide some concrete and illustrative examples of how computational approaches, including the use of LLMs, can be useful in legal scholarship, drawing on my own and others' research.

3. USEFULNESS OF COMPUTATIONAL METHODS IN LEGAL SCHOLARSHIP: SOME EXAMPLES

3.1 Generative AI to the Rescue?

Computational methods are not new, not even new in the field of law. Researchers in the field of legal informatics, as well as commercial actors in the Legal Information Retrieval Systems (LIRS) market and the so-called LegalTech sector, have been applying computational methods to law for some time.[31] As I will elaborate on, these methods have also been used in legal scholarship for quite some time.

The introduction of GPT-3 in 2020 likely provided many lawyers' first direct experience with the power of applying computational methods to law.[32] ChatGPT and other chatbots that are based on generative, pre-trained large language models (LLMs) have proved capable of accurately answering quite sophisticated questions about the law.[33] It is true that even state-of-the-art LLM-based chatbots specifically fine-tuned on legal data are prone

3.2 A Helicopter Perspective on the Law

Some of the most important contributions that conducting computational and other large-n studies can provide come from the type of questions that they allow us to ask. While these benefits may come across as somewhat "soft", they should not be underestimated. My first experience with using computational methods was born out of a dissatisfaction over the natural, cognitive limitations on the size of the dataset that one can analyze using a purely doctrinal approach. My dissertation had left me with the impression that the Court of Justice of the European Union (CJEU) was inconsistent in how it cited and used its own case law,[37] but to identify the existence and absence of citation patterns on a large scale was impossible using traditional legal methods. However, by using network analysis we were able to study all references in and between all decisions. I have since come to appreciate that just as some important aspects of the law can only be understood through a close reading, others, like the Nazca Lines of Peru, only make sense when viewed from high above.

Doctrinal legal research largely rests on deductive reasoning, that is to say that it departs from predefined prin-

---

31   Margoni (n 14) 490–493.

32   Not every reader may be aware that the use of pre-trained LLMs in law predates GPT-3. See e.g. Ilias Chalkidis and others, 'LEGAL-BERT: The Muppets Straight Out of Law School', *Findings of EMNLP* (Association for Computational Linguistics 2020).<https://www.aclweb.org/anthology/2020.findings-emnlp.261> accessed 17 March 2023).

33   Daniel Martin Katz and others, 'GPT-4 Passes the Bar Exam' [2023] SSRN Electronic Journal; Jonathan H Choi and others, 'ChatGPT Goes to Law School' [2023] SSRN Electronic Journal <https://www.ssrn.com/abstract=4335905> accessed 17 March 2023; Michael James Bommarito and Daniel Martin Katz, 'GPT Takes the Bar Exam' [2022] SSRN Electronic Journal.

34   Varun Magesh and others, 'Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools'.

35   Andrew Blair-Stanek, Nils Holzenberger and Benjamin Van Durme, 'Can GPT-3 Perform Statutory Reasoning?' <http://arxiv.org/abs/2302.06100> accessed 4 February 2024.

36   Their current level of competence can perhaps be compared to that of an experienced law student or recent law graduate.

37   Johan Lindholm, State Procedure and Union Rights: A Comparison of the European Union and the United States (Iustus Förlag 2007).

ciples, concepts, and rules. In this regard computational approaches enhance our ability to conduct inductive legal research by identifying patterns in very large legal datasets.[38] Of particular interest in this regard are so-called unsupervised approaches, that is to say approaches that allow computers to identify patterns unrestrained from any preconceived notions or theories. By allowing computers to "run free" they can identify patterns in empirical data that challenges lawyers' existing theories about the law. For example, using network analysis, we clustered all of the CJEU's decisions into communities based on how they are connected to each other through citations. Those communities are functionally comparable to areas of law, but because we used an unsupervised approach they could and in some important regards did differ from the areas of EU law one encounters in textbooks.[39] Citations are not the only method that can be used for categorizing legal sources in an unsupervised manner. A commonly used computational approach is to identify topics in large legal text collections and to categorize individual legal text on these topics using topic modelling.[40] This has a number of useful applications. For example, I have used unsupervised topic modelling to split sports arbitration cases into novel categories,[41] whereas Yannis Panagis, Martin Lolle Christensen and Urška Šadl[42] used it to track legal change in European courts.

This is also an illustration of how computational methods can assist in theorizing about law.[43] It is important to point out that machine-identified patterns are not absolute truths and that they should always be subject to human analysis. However, if the facts do not seem to fit the theory, it is good scientific practice to change the theory, and computational approaches can help us test to what extent theories and facts fit together. Also, conducting empirical research can help us sharpen legal theories and concepts. A theory can only be tested and a concept captured or quantified if it is clear. By requiring sharp legal theories and concepts, the use of computational

approaches reveal ambiguities and inconsistencies. As Zglinski acutely observes, "[a]nalysing large numbers of decisions forces us to be precise about what it is that we are looking for [and i]t, thereby, indirectly benefits the conceptual work."[44] For example, in one study, we used network analysis to identify which CJEU decisions are the most "important". This required us to clarify the different ways in which a case can be legally important and turn them into measurable variables.[45] Similarly, in his study of judicial deference, Zglinski developed a conceptual framework in order to study its presence in the CJEU over time.[46]

3.3 Generating Research Data

Empirical research is only as good as the data that is based on. While legal scholars are very skilled at collecting, systematizing, and analyzing legal authorities, the methods traditionally used in law are poorly suited for generating accurate and reproducible large legal datasets that capture relevant internal aspects of the law.[47] A fundamental challenge in this regard is that the variables that we want to capture are hidden in complicated, technical, and nuanced language found in documents that are, at best, semi-structured.[48] For example, a legal scholar or a legally-trained research assistant reading CJEU decisions can determine whether the Court deferred to a national court to make the final decision, but to do this on a large scale is prohibitively time consuming.[49] What if we can train computers to be perfect research assistants: perfectly consistent, highly effective, low-cost, and able to work indefinitely without taking breaks?[50]

Significant progress towards this becoming reality has been made in recent years. One example of this that should interest legal scholars is Joana Ribeiro De Faria, Huiyuan Xie and Felix Steffek[51] who successfully employed GPT-4 to extract key legal aspects from case law text, such as claims, references, case outcomes, and

38   Margoni (n 14) 491–492.

39   Atieh Mirshahvalad and others, 'Significant Communities in Large Sparse Networks' (2012) 7 PLoS ONE e33721; Mattias Derlén and others, 'Coherence Out of Chaos: Mapping European Union Law by Running Randomly Through the Maze of CJEU Case Law' (2013) 16 Europarättslig Tidskrift 517; see also Martin Lolle Christensen, Henrik Palmer Olsen and Fabien Tarissan, 'Identification of Case Content with Quantitative Network Analysis: An Example from the ECtHR', vols 29th International Conference on Legal Knowledge and Information Systems (JURIX'16) (2016) <https://hal.science/hal-01386810> accessed 20 August 2024.

40   See e.g. M Mohammadi and others, 'Combining Topic Modelling and Citation Network Analysis to Study Case Law from the European Court on Human Rights on the Right to Respect for Private and Family Life' <http://arxiv.org/abs/2401.16429> accessed 14 August 2024; Tan and others (n 11).

41   Johan Lindholm, 'Court of Arbitration for Sport: En framgångsrik trettiofemåring med begynnande medelålderskris?' (2019) 2019/20 Juridisk Tidskrift 482, 146–159.

42   'On Top of Topics: Leveraging Topic Modeling to Study the Dynamic Case-Law of International Courts' (2016) 294 Frontiers in Artificial Intelligence and Applications 161.

43   Cf. Michael Heise, 'The Importance of Being Empirical' (1999) 26 Pepperdine Law Review 807, 813 ("The development of good theories is made even more difficult without the benefit of good data.").

44   Jan Zglinski, Europe's Passive Virtues: Deference to National Authorities in EU Free Movement Law (Oxford University Press, Oxford 2020), 7.

45   Mattias Derlén and Johan Lindholm, 'Goodbye van Gend En Loos, Hello Bosman? Using Network Analysis to Measure the Importance of Individual CJEU Judgments' (2014) 20 European Law Journal 667.

46   Ibid. See also Michal Ovádek, Phillip Schroeder and Jan Zglinski, 'Where law meets data: a practical guide to expert coding in legal research' European Law Open (forthcoming); Jan Zglinski, 'What is the Point of Empirical Legal Research in EU Law?' in Empirical Legal Studies in EU Law (Cambridge University Press, Cambridge, forthcoming).

47   Cf. e.g. Mark A Hall and Ronald F Wright, 'Systematic Content Analysis of Judicial Opinions' (2008) 96 California Law Review 63; Frankenreiter and Livermore (n 12) 40.

48   The task of extracting the valuable information is sometimes referred to as text and data mining or TDM. See e.g. Margoni (n 14) 487.

49   The coding task that is ideal for automation is one that is sufficiently clear that humans can reliably do it but it takes a lot of time.

50   Cf. Alessandro Contini and others, 'Recognising Legal Characteristics of the Judgments of the European Court of Justice: Difficult but Not Impossible' [2022] Legal Knowledge and Information Systems.

51   'Automatic Information Extraction from Employment Tribunal Judgements Using Large Language Models' [2024] SSRN Electronic Journal <https://www.ssrn.com/abstract=4776160> accessed 4 June 2024.

reasons for the decision. Ivan Habernal and others[52] were similarly able to use LLMs to 'mine' different types of legal arguments, such as different methods of interpretation, in the case law of the European Court of Human Rights. A third and final example is Jonathan H Choi who developed a computational method for measuring the clarity of legal texts.[53]

Named entity recognition (NER) is a computational task that should be of great interest to legal scholars. NER involves the identification of unique identifiers of 'entities' in text, such as proper nouns, names referring to people or places, but which in principle can be any type of text element.[54] Legal texts are full of entities that are of central relevance when it comes to understanding the text, including legal-internal entities such as sources, actors, rules, principles, and legal concepts. I would think that the ability to reliably and effectively identify such legal entities in very large amounts of legal text makes NER valuable to most legal scholars.[55] Scholars have been developing methods for NER in legal text and successfully applied these to extract a variety of legal contexts across multiple jurisdictions.[56] Being able to annotate references to legal concepts in legal text automatically, reliably, and on a large scale creates a number of opportunities for legal scholars. In addition to the value of information about legal entities as such, using NER-annotated text data can enhance other computational methods in law.[57]

Another example of how computational methods can be useful in generating valuable legal research data involves 'issue splitting'. It is not uncommon that judgments address multiple, distinct legal issues and for each such issue contain reasoning and holding. This makes entire judgments a non-ideal unit of observation for the purpose of empirically studying case law.[58] Judgments often contain extraneous information, such as details about the parties, quotes from relevant legislation, and costs, that may not only be irrelevant but that for analytical purposes constitutes "noise" and that ideally should be removed. Paragraphs or sentences on the other hand are too fine of a unit as important contextual information is lost. Schroeder and I therefore propose the concept of legal issues as a 'Goldilocks' layer, a more efficient level of analysis that balances comprehensiveness and specificity. While it is possible to hand-split judgments by legal issues, it is an immensely resource-intensive task. As an alternative, we hand-coded a relatively small set of judgments by the CJEU and used this data to train a neural network to identify where the Court starts and stops discussing a legal issue. We then use this model to quickly, cheaply, and with high accuracy 'issue split' a much larger number of CJEU decisions.[59]

These examples of successful automated coding of legal data – using computers as research assistants – worked well because the coding tasks were relatively easy, that is to say that the concepts of interest were clear and rather simple and that they were expressed in the text in a transparent and reliable fashion. This will not always be the case. In fact, scholars are often most interested in complex and vague concepts that are difficult to reliably code even by hand and after much training. Although advances in machine learning techniques constantly moves the frontier forwards, some tasks will be beyond machines' reach for a long time (and forever unless scholars do the necessary conceptualization and theorization). Humans and machines are good at and should be used for different tasks: whereas machines are ideal research assistants solving many tedious tasks, the hardest problems should be left to humans.[60]

3.4 Predicting Citations

A critical aspect of legal research and practice involves identifying relevant sources, such as case law, that support legal propositions. Given the rapidly expanding volume of legal documents – such as the more than 800 judgments issued by the CJEU each year – this task is becoming increasingly difficult to perform effectively without computer assistance, and eventually possibly impossible.[61] This raises the question: can we predict citations for legal propositions at a paragraph level? To address this, a group of scholars that included myself attempted to mimic CJEU citation patterns by estimating the probability that the CJEU would cite a particular paragraph in support of a legal statement, based on previous citations. Our approach involves training a BERT-

52   Ivan Habernal and others, 'Mining Legal Arguments in Court Decisions' [2023] Artificial Intelligence and Law.

53   'Measuring Clarity in Legal Text' (2024) 91 The University of Chicago Law Review 1.

54   Mónica Marrero and others, 'Named Entity Recognition: Fallacies, Challenges and Opportunities' (2013) 35 Computer Standards & Interfaces 482.

55   Cf. Christopher Dozier and others, 'Named Entity Recognition and Resolution in Legal Text' in David Hutchison and others (eds), *Semantic Processing of Legal Texts*, vol 6036 (Springer Berlin Heidelberg 2010) 1–3 <http://link.springer.com/10.1007/978-3-642-12837-0_2> accessed 26 June 2024.

56   See e.g. Elena Leitner, Georg Rehm and Julián Moreno-Schneider, 'Fine-Grained Named Entity Recognition in Legal Documents' [2019] SEMANTiCS 2019 272; Vitor Oliveira and others, 'Combining Prompt-Based Language Models and Weak Supervision for Labeling Named Entity Recognition on Legal Documents' [2024] Artificial Intelligence and Law; Andreas Östling and others, 'The Cambridge Law Corpus: A Corpus for Legal AI Research' <http://arxiv.org/abs/2309.12269> accessed 22 September 2023; Milagro Teruel and others, 'Legal Text Processing Within the MIREL Project' in Georg Rehm, Victor Rodríguez-Doncel and Julián Moreno-Schneider (eds) (2018); Ilias Chalkidis, Ion Androutsopoulos and Achilleas Michos, 'Extracting Contract Elements', *Proceedings of ICAIL '17* (2017).

57   Irene Benedetto and others, 'Boosting Court Judgment Prediction and Explanation Using Legal Entities' [2024] Artificial Intelligence and Law.

58   E.g. when answering research questions about courts and law or for offering well-informed recommendations.

59   Philipp Schroeder and Johan Lindholm, 'From One to Many: Identifying Issues in CJEU Jurisprudence' (2023) 11 Journal of Law and Courts 163.

60   See also, for a similar argument in math, interview with Terrence Tao in Matteo Wong, "We're Entering Uncharted Territory for Math", The Atlantic, 4 October 2024.

61   Cf. Benjamin Alarie, 'The Path of the Law: Towards Legal Singularity' (2016) 66 University of Toronto Law Journal 443; Simon Deakin and Christopher Markou, 'From Rule of Law to Legal Singularity' in Simon Deakin and Christopher Markou (eds), *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence* (Hart Publishing 2020).

based encoder model on both positive (cited) and negative (not cited) text data to predict citation links between paragraphs. Our model, when tested, on average ranks the actually cited paragraph as number 2. This method enables us to predict references, identify surprising references, and model relationships between legal statements and their supporting sources. Ultimately, this approach provides valuable tools for offering recommendations, as well as for detecting, studying, and explaining unexpected judicial reasoning.[62]

An accurate citation prediction model has multiple potential uses in legal scholarship. One of these is to measure whether a decision is 'good law', that is to say whether it is a good authority for a legal proposition, or whether it for example has been overruled or become obsolete. To do so has important legal implications and tangible practical uses. For example, while explicit overruling is rare on the CJEU, the Court frequently implicitly or 'covertly' overrules its own case law.[63] Consequently, it can be difficult to know whether CJEU case law is good law. Hand coding whether a case is good law is possible,[64] but to do so is prohibitively expensive. This begs the question, can 'case law health' be measured computationally? We have previously experimented with network analysis, more specifically various network centrality measurements, to capture whether a CJEU decision has been subsequently overruled.[65] Access to a model capable of accurately predicting the probability that a judgment would be cited in support in a particular textual context and being able to measure the difference between a case's predicted and observed citation rate provides an exciting new avenue in this field.

4. WHAT DO WE NEED GOING FORWARD?

I hope that I have convinced the reader that it is both appropriate and useful to employ computational methods in legal research, and that Chau and Livermore are correct in that computational methods, "[u]sed in conjunction with traditional legal research methodologies,… promise to open new avenues of research that could revolutionize the study of law."[66] It seems to me that state-of-the-art computational approaches are ideally suited for law that rule-based approaches, due to law's indeterminate features, fails to capture accurately and fully.

My position on this matter is supported by an arguably liberal understanding of legal scholarsip. The core mission of social sciences and scientists is to produce novel insights about social phenomena. In the specific case of

---

62   Henrik Palmer Olsen and others, 'Re-Framing Case Law Citation Prediction from a Paragraph Perspective' in Giovanni Sileno, Jerry Spanakis and Gijs Van Dijck (eds), *Legal Knowledge and Information Systems* (IOS Press 2023).

63   Daniel Sarmiento, 'The 'Overruling Technique' at the Court of Justice of the European Union' [2023] European Journal of Legal Studies 109; Jan Komárek, 'Judicial Lawmaking and Precedent in Supreme Courts' [2011] SSRN Electronic Journal 32–33 <http://www.ssrn.com/abstract=1793219> accessed 6 May 2022.

64   Some American actors in the LIRS space provide this service.

65   Mattias Derlén and Johan Lindholm, 'Is It Good Law? Network Analysis and the CJEU's Internal Market Jurisprudence' (2017) 20 Journal of International Economic Law 257.

66   Chau and Livermore (n 12) 10.

legal scholarship and scholars, that social phenomenon is the law. Whereas producing new knowledge about the law requires the scientific community to employ certain methods,[67] it does not, on principle or in practice, exclude other methods. Every method is obviously not a good fit for answering every research question, but methodological conservatism also has no value *per se*. Moreover, I hope that I, through my examples, have been able to convince some readers that computational methods can and have helped produce novel insights into law, even from a legal-internal perspective.

Frankenreiter and Livermore write that "[a]s these tools continue to advance, and law scholars become more familiar with their potential applications, the impact of computational methods is likely to continue to grow."[68] While I hope that this will be true, I am somewhat sceptical about the ease of the transition. In order to advance the use of computational methods in legal research we must, first, improve access to the infrastructure on which it is based. Legal data access has improved significantly in recent years, but access to open, reliable, and comprehensive legal datasets still constitutes a bottleneck. Such data needs to include not only statutes and court precedent, but also, *inter alia*, preparatory works, other documents from the legislative process, legal literature, decisions by lower courts and administrative agencies, and party court filings. Two major obstacles to the development of such datasets is that not all legal sources are collected in a freely and publicly accessible archive and that access to important information about law and legal institutions are blocked by commercial actors that hold intellectual property rights. It is however not sufficient to make text data available, one must also ensure that it is accurate and of high quality. This means ensuring that legal texts are curated, clean, correct, accurate, and organized. Additionally, it is essential to enhance it with rich and accurate metadata, including the use of unique and stable identifiers, assigning legally relevant labels to text elements, and tagging of natural and legal entities. Achieving this

requires the collaboration of multiple actors including libraries, parliament, government, courts, government agencies, and commercial actors. On the academic side, the creation of shared and open legal datasets will require pooling the skills and efforts of legal scholars, computer scientists, and other academics.

Law schools and legal scholars also have an important role to play when it comes to capacity building. Currently, most European law students graduate without serious exposure to empirical methods or research design, creating a "closed loop" from which professors and doctoral students are drawn, perpetuating stagnant methodological capacities. To break this loop, it is essential to introduce doctoral students to empirical legal thinking and computational methods, thus fostering a new generation of scholars equipped with the tools necessary for modern legal research. Not every future legal scholar will or should learn to master state-of-the-art computational methods, but if we can provide them with a basic understanding of the tools, their possibilities, and their limitations, we can facilitate fruitful collaboration between legal scholars and computer scientists.[69]

Johan Lindholm

Professor of Law, Umeå University, Department of Law. email: johan.lindholm@umu.se. This essay is based on talks that he gave at the Conference on Text and Data Mining and Artificial Intelligence, 14 June 2024, at Stockholm University, and at a seminar at the Department of Law at Uppsala University. The author is grateful for comments and suggestions provided by the participants, as well as by Isak Nilsson and Jan Zglinski.

---

67 Most obviously to engage in traditional descriptive and normative jurisprudence.

68 Frankenreiter and Livermore (n 12) 39.

69 Kantorowicz-Reznichenko (n 6) 5; Heise (n 4) 828–829.