

# Computational Methods for Legal Analysis

## The Way Forward?

Elena Kantorowicz-Reznichenko\*

### Abstract

Computational analysis can be seen as the most recent innovation in the field of Empirical Legal Studies (ELS). It concerns the use of computer science and big data tools to collect, analyse and understand the large and unstructured data, such as for instance (legal) text. Given that the text is now the object of analysis, but the methods are (largely) quantitative, it lies in the intersection between doctrinal analysis and ELS. It brings with it not only a great potential to scale up research and answer old research questions, but also to reveal uncovered patterns and address new questions. Despite a slowly growing number of legal scholars who are already applying such methods, it is underutilised in the field of law. Furthermore, given that this method comes from social and computer sciences, many legal scholars are not even aware of its existence and potential. Therefore, the purpose of this special issue is not only to introduce these methods to lawyers and discuss possibilities of their application, but also to pay special attention to the challenges, with a specific emphasis on the ethical issues arising from using 'big data' and the challenge of building capacity to use such methods in law schools. This editorial briefly explains some of the methods which belong to the new movement of Computational Legal Analysis and provides examples of their application. It then introduces those articles included in this special issue. Finally, it provides a personal note on the way forward for lawyers within the movement of Computational Legal Analysis

**Keywords:** computational legal analysis, empirical legal studies, natural language processing, machine learning

## 1 Introduction

Traditional doctrinal analysis is the backbone of legal scholars. It entails reading legal documents such as court decisions and legislations, and then offering descriptive, critical, normative or predictive claims about different legal rules or decisions.<sup>1</sup> Empirical Legal Studies (ELS), and especially quantitative ELS, use

statistical tools to analyse legal phenomena/issues. This analysis does not need to be performed on legal text. For example, legal questions can be answered by conducting experiments in the lab, online or in the field where a legal situation is simulated, and the actions or choices of actors are recorded,<sup>2</sup> or it can involve vignette studies of legal practitioners, for instance, examining whether cognitive biases affect judges' decisions.<sup>3</sup> ELS also involve the analysis of observational data, using different natural experiment research designs to be able to infer causality, for example, measuring whether an increase in the numbers of police officers enhances deterrence of crimes.<sup>4</sup> In recent decades, unprecedented technological advances in artificial intelligence (AI) tools and an ever-increasing digitalisation of legal documents can be witnessed. This combination has led to the emergence of a new research stream in law – *Computational Legal Analysis (CLA)*. This type of research lies at the intersection of doctrinal analysis and quantitative ELS. The focus of CLA is the legal text, the same as traditional doctrinal analysis, but it uses computer science and statistical tools to collect, analyse and understand the texts, which are now treated as empirical data.<sup>5</sup> Some of these methods allow the scaling up of the research that has been conducted until now by lawyers using, for example, manual coding of texts. Other methods, as discussed in the next section, make it possible to answer new questions and uncover patterns in legal documents or links between them, which are not easily detectable through hand-coded analysis.

\* Elena Kantorowicz-Reznichenko is Professor of Quantitative Empirical Legal Studies at the Rotterdam Institute of Law and Economics, Erasmus School of Law, Erasmus University, Rotterdam. I would like to thank Nina Holvast, Jaroslaw Kantorowicz and Pim Jansen for their useful comments, and Vera Brijer for her editorial assistance.

1. T. Hutchinson, 'The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law', 8 *Erasmus Law Review* 130 (2015).

2. See for example, C. Engel and M. Kurschilgen, 'Fairness Ex Ante and Ex Post: Experimentally Testing Ex Post Judicial Intervention into Blockbuster Deals', 8 *Journal of Empirical Legal Studies* 682 (2011); E. Kantorowicz-Reznichenko, J. Kantorowicz & K. Weinsall, 'Can We Overcome Ideological Biases in Constitutional Judgments? An Experimental Analysis', *DIIS Working Paper*, [https://www.researchgate.net/publication/353333715\\_Can\\_We\\_Overcome\\_Ideological\\_Biases\\_in\\_Constitutional\\_Judgments\\_An\\_Experimental\\_Analysis](https://www.researchgate.net/publication/353333715_Can_We_Overcome_Ideological_Biases_in_Constitutional_Judgments_An_Experimental_Analysis).
3. C. Guthrie, J.J. Rachlinski & A.J. Wistrich, 'Inside the Judicial Mind', 86 *Cornell Law Review* 777-830, 788-791 (2000); C. Guthrie, J.J. Rachlinski & A.J. Wistrich, 'The 'Hidden Judiciary': An Empirical Examination of Executive Branch Justice', 58 *Duke Law Journal* 1477, at 1502-4 (2009).
4. R. Di Tella and E. Schargrodsky, 'Do Police Reduce Crime? Estimates Using the Allocation of Police Forces after a Terrorist Attack', 94 *American Economic Review* 115-33 (2004); J. Klick and A. Tabarrok, 'Using Terror Alert Levels to Estimate the Effect of Police on Crime', 48 *Journal of Law & Economics* 267-79 (2005). For ELS in Europe and also more information on qualitative methods used in ELS, see the special issue by P. Mascini and W. van Rossum, 'Empirical Legal Research: Fad, Feud or Fellowship?', *Erasmus Law Review* 2 (2018).
5. M.A. Livermore and D.N. Rockmore, 'Law as Data: Computation, Text, & the Future of Legal Analysis', *Santa Fe Institute Press* xvii (2009).

The purpose of this special issue is to raise the awareness of lawyers and legal scholars to the existence of these new methods. It is meant to present the promises of CLA and its potential uses, its challenges, including ethical ones, and some thoughts on the training of law students and legal scholars in these methods. The starting point is that ELS in general is an important field, and my arguments to support this point can be found in my previous work.<sup>6</sup> The goal of this special issue is to focus on CLA, which I consider as an important innovation and addition to ELS. In this editorial, I first briefly explain which type of research and practical application is possible using computational methods. Next, I introduce the articles in this special issue and explain how they connect to the story of CLA. I end with a personal note on the future of CLA and legal education.

## 2 What Are CLA Methods, and Which Type of Research Can They Support?

CLA concerns the application of computing capabilities to law as a subject matter.<sup>7</sup> It includes different techniques such as network analysis, machine learning and natural language processing (NLP).<sup>8</sup> Utilising computational power allows researchers to investigate simpler things such as the features of the legal text (for example, length of judicial opinions, linguistic sophistication of judicial opinions) or the prominence of certain decisions by creating a citation network and exploring which cases are the most cited. However, more complicated research questions can also be answered using CLA methods.<sup>9</sup> Computational methods can also be used to simply investigate the evolution of a research field itself<sup>10</sup> or to synthesise fields of research.<sup>11</sup> Here I mention several studies as examples, without entering into the details, to briefly provide a sense of the potential of these methods. A more thorough review of some of these methods – in the context of quantitative text analysis – can then be

found in the first article in this special issue by Arthur Dyevre.

One interesting stream of research utilises plagiarism-detection software, or related techniques, which checks for unusual similarities between texts, to investigate the influence of other (legal) texts or people on judicial decisions. For example, different studies use such techniques to identify whether judges write their decisions themselves or whether, more often, they are assisted by clerks;<sup>12</sup> which countries have a stronger influence on WTO decisions;<sup>13</sup> whether texts from international treaties are being copy-pasted into new international agreements;<sup>14</sup> or to what extent Supreme Court judges are influenced by and use the arguments of lower courts.<sup>15</sup> These techniques are useful for lawyers and legal scholars interested in the impact that different actors have on judicial and legislative decisions. A sensational example of the relevance of such an analysis to law can be found in the contested arbitration award imposed on the Russian Federation in the case *Yukos v Russia*.<sup>16</sup> Appealing the award in Dutch courts, Russia bought an expert report in which syntactic analysis was used to compare the award decision against the previous writings of the arbitrators on the one hand, and the Tribunal's assistant on the other hand. The linguistic analysis demonstrated that substantial parts of the Tribunal's decision were written by the assistant rather than the arbitrators.<sup>17</sup> Such methods can also be used to examine which texts and groups influenced legislation, for example, whether and how reports of lobby groups are incorporated into the draft legislation.<sup>18</sup>

Another interesting application of computational methods in law is the citation network analysis. It assists lawyers, among others, in identifying landmark cases, detecting trends of precedents' importance over time and exploring potentially overlooked precedents.<sup>19</sup> For example, one study was seeking to identify the most

6. E. Kantorowicz-Reznichenko, 'Lawyer 2.0! Some Thoughts on the Future of Empirical Legal Studies in Europe' in R. van den Bergh, M. Faure, W. Schreuders & L. Visscher (eds.), *Don't Take it Serious: Essays in Law and Economics in Honour of Intersentia* (2018).
7. See, for example, the definition of Computational Legal Studies in R. Whalen, 'The Emergence of Computational Legal Studies: An Introduction', in R. Whalen (ed.) *Computational Legal Studies: The Promise and Challenge of Data-Driven Research*, Edward Elgar Publishing Limited (2020) 1-8 at 2.
8. J. Frankenreiter and M.A. Livermore, 'Computational Methods in Legal Analysis', 16 *Annual Review of Law and Social Science* 39, at 21.2 (2020).
9. *Ibid.*, at 21.5
10. E. Kantorowicz-Reznichenko and J. Kantorowicz, 'Law & Economics at Sixty – Mapping the Field with Bibliometric and Machine Learning Tools', *DIIS Working Paper* 2020.
11. S. Kuipers, J. Kantorowicz & J. Mostert, 'Manual or Machine? A Review of the Crisis and Disaster Literature', 10 *Risk, Hazards & Crisis in Public Policy* 4, 388-402 (2019) doi: <https://doi.org/10.1002/rhc3.12181>.

12. S.J. Choi and G.M. Gulati, 'Which Judges Write Their Opinions (and Should We Care)', 32 *Florida State University Law Review* 1077 (2004).
13. M. Daku and K.J. Pelc, 'Who Holds Influence Over WTO Jurisprudence?', 20 *Journal of International Economic Law* 233-55 (2017).
14. T. Allee and M. Elsig, 'Are the Contents of International Treaties Copied and Pasted? Evidence from Preferential Trade Agreements', 63 *International Studies Quarterly* 603-13 (2019).
15. P.C. Corley, P.M. Collins Jr & B. Calvin, 'Lower Court Influence on US Supreme Court Opinion Content', 73 *The Journal of Politics* 31-44 (2011).
16. PCA Case No. AA227, *Yukos Universal Limited (Isle of Man) v. Russia*.
17. J. Hepbur, 'Battling \$50 Billion Yukos Awards On Two Fronts, Russia Focuses On Claimants' Alleged Fraud And Linguistic Analysis Of Tribunal Assistant's Alleged Role In Drafting Awards, Investment Arbitration Reporter', (2015), [www.iareporter.com/articles/battling-50-billion-yukos-awards-on-two-fronts-russia-focuses-on-claimants-alleged-fraud-and-linguistic-analysis-of-tribunal-assistants-alleged-role-in-drafting-awards/](http://www.iareporter.com/articles/battling-50-billion-yukos-awards-on-two-fronts-russia-focuses-on-claimants-alleged-fraud-and-linguistic-analysis-of-tribunal-assistants-alleged-role-in-drafting-awards/) (last visited 7 July 2021).
18. M. Burgess, E. Giraudy, J. Katz-Samuels, J. Walsh, D. Willis, L. Haynes & R. Ghani, 'The Legislative Influence Detector: Finding Text Reuse in State Legislation', *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 57-66 (2016).
19. D. van Kuppevelt, G. van Dijk & M. Schaper, 'Purposes and Challenges of Legal Citation Network Analysis on Case Law', in R. Whalen (ed.) *Computational Legal Studies: The Promise and Challenge of Data-Driven Research*, Edward Elgar Publishing Limited (2020) 265.

important cases of the Court of Justice of the European Union (CJEU). By ‘importance’, the authors meant those cases that are the most cited and applied to resolve disputes at hand. To achieve this, the authors used a ‘network analysis’, where they included all the cases of the CJEU since the beginning of its functioning. The value of this analysis was that they have shown that the cases that are normally assumed to be the most important jurisprudence of the CJEU are not actually the most important in terms of constituting precedents which are applied frequently in later cases.<sup>20</sup> A similar exercise was conducted in respect of other courts. For example, another study investigated the impact of different cases of the European Court of Human Rights (ECtHR) in terms of frequency of citations. The underlying assumption was that these are the cases that can be considered as precedents for general principles of law.<sup>21</sup> The final application of computational methods for lawyers which I would like to present here are predictive models. It is now possible, and this is increasingly being done, to train a machine learning model, an algorithm, to predict the outcome of specific cases. To explain the method in a very simple way, a dataset, which, for example, consists of court decisions, is randomly divided into a training set and a testing set. The first subset of the data is used to train the algorithm to recognise certain patterns. Once the model is built, it is tested on the other subset of the data to see whether it can predict the outcome in those documents. When the model is sufficiently accurate, it can be used to predict the outcomes of cases that are yet to be resolved (out-of-sample data set).<sup>22</sup> For example, in one study, political scientists who used a statistical model were compared to 83 legal experts in their ability to predict the outcome of upcoming American Supreme Court cases accurately. The outcome was that the statistical model had an accuracy rate of 75% of the cases whereas the legal experts were right 59% of the time.<sup>23</sup> Attempts to predict decisions of courts were also made in Europe, for example, the decisions of the ECtHR.<sup>24</sup>

From the description above, it is clear that predictive models are not only useful for researchers, but also, and

maybe even more so, for legal practitioners. Some examples are discussed in detail by Simon Vydra and co-authors in the second article in this issue. Here I would like to briefly present one example. The criminal justice system often involves predictions. For example, the police need to predict in which areas to concentrate their efforts, judges need to predict risk levels of offenders when deciding on the type of arrest (home or jail) and parole decisions heavily depend on the expected level of risk of the convicted offender. Therefore, machine learning–assisted prediction has been found useful in this field.<sup>25</sup> For example, a machine learning model was used in bail decisions in the belief that it can reduce the rate of imprisonment without increasing the risk of crime.<sup>26</sup>

After briefly presenting CLA’s potential, I would like to point out one important limitation.<sup>27</sup> Especially with predictive models, one could be tempted to interpret prediction results as offering causal links. With traditional empirical methods, we are usually interested in investigating how A causes B. For example, we seek to understand how the new directive on digital copyright will affect creators’ incentives when contracting their copyrights, or how changing employees’ protection rules will affect the flexibility of the labour market and the behaviour of employers. Computational methods, on the other hand, are very effective in providing prediction of outcomes. As I have shown in the previous section, by utilising large data on past behaviour we can predict future behaviour. However, this does not enable us to immediately refer to causal links between different factors. In other words, the fact that, for example, some combination of words or other factors predicts a certain outcome does not mean that these factors *cause* this outcome.<sup>28</sup> One illustrative example can be found in the above-mentioned study on ECtHR decisions. Among the words/combinations of words that predicted a decision by the court that a violation took place was the date October 2007.<sup>29</sup> Clearly, there is no causal link between the date and the question of whether there was violation or not. Because machine learning techniques are not normally suitable for inference questions, to fully utilise their potential, researchers should identify those questions that can be meaningfully answered by prediction and classification.<sup>30</sup> Having said that, one should note

20. M. Derlén and J. Lindholm, ‘Goodbye van Gend en Loos, Hello Bosman? Using Network Analysis to Measure the Importance of Individual CJEU Judgments’, 20 *European Law Journal* 667-87 (2014).
21. H.P. Olsen and M. Esmark, ‘Needles in a Haystack: Using Network Analysis to Identify Cases that are Cited for General Principles of Law by the European Court of Human Rights’, in R. Whalen (ed.) *Computational Legal Studies: The Promise and Challenge of Data-Driven Research*, Edward Elgar Publishing Limited (2020) 293-311.
22. R. Copus, R. Hübert & H. Laqueur, ‘Big Data, Machine Learning, and the Credibility Revolution in Empirical Legal Studies’, in Michael A. Livermore and Daniel N. Rockmore (eds.) *Law as Data: Computation, Text & the Future of Legal Analysis*, Santa Fe: Santa Fe Institute of Science (2019) 21-37, at 31-3.
23. A.D. Martin, K.M. Quinn, T.W. Ruger & P.T. Kim, ‘Competing Approaches to Predicting Supreme Court Decision Making’, 2 *Perspectives on Politics* 4, 761-767 (2004).
24. M. Medvedeva, M. Vols & M. Wieling, ‘Using Machine Learning to Predict Decisions of the European Court of Human Rights’, 28 *Artificial Intelligence and Law* 237-66 (2020). (The prediction is within their data set, and not necessarily already advanced enough to forecast a future decision, even though that is the eventual goal with such methods.)

25. Copus et al., above n. 22, at 31-3.
26. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig & S. Mullainathan, ‘Human Decisions and Machine Predictions’, 133 *The Quarterly Journal of Economics* 237-93 (2018).
27. This is not to say there are no other limitations, and some will be covered in the second article in this special issue, which discusses the ethical side of using computational methods.
28. M. Dumas and J. Frankenreiter, ‘Text as Observational Data’, Michael A. Livermore and Daniel N. Rockmore (eds.) *Law as Data: Computation, Text & the Future of Legal Analysis*, Santa Fe: Santa Fe Institute of Science (2019) 59-70, at 59-65.
29. Presentation of the paper M. Vols, ‘Using machine learning to predict decisions of the European Court of Human Rights’, ATLAS AGORA Summer School, Erasmus School of Law, Erasmus University Rotterdam (25 June 2021).
30. Dumas and Frankenreiter, above n. 28, at 63.

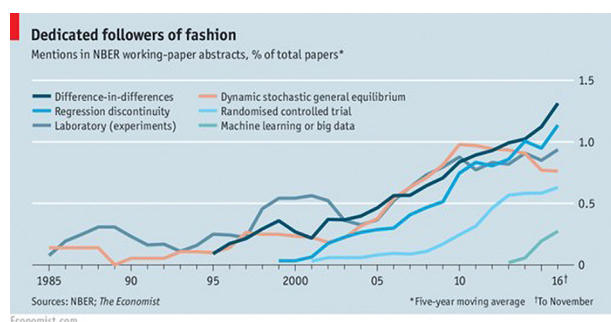
that the stream of causal machine learning is gaining more and more traction in academic literature.<sup>31</sup>

Predictions are easier to do than causal inference. And sometimes we would like to predict something rather than insisting on understanding the causes (for instance, which type of offenders are expected to reoffend, without identifying the exact causes, allows us to focus resources on such offenders).<sup>32</sup> Furthermore, given the complexity of causal inference, it might be suggested to first use a prediction model in order to establish that a certain variable can predict changes in the other, and only then turn one's attention into the work of causal inference.<sup>33</sup>

### 3 What Is This Special Issue About?

Computational analysis as such is not new. It has already been implemented in other fields like social sciences and digital humanities. See, for example, Figure 1 for the increase in the use of machine learning in economics research.

Figure 1 The Use of Different Quantitative Methods by Economists



Source: Economists Are Prone to Fads, and the Latest Is Machine Learning: Big Data Have Led to the Latest Craze in Economic Research, *The Economist*, 24 November 2016.

However, it is a relatively new movement for the legal world and seems for now to be overlooked by legal scholars, especially in Europe. Despite a slowly growing group of legal computational scholars, the number of legal scholars who implement those methods in their research, or are even aware of their existence, is still limited. Therefore, the main purpose of this special issue is to simply open up the discussion, and to raise awareness of CLA among lawyers in Europe and to invoke their curiosity. Most lawyers and legal scholars are not skilled in computational methods. However,

31. See for example, H. Farbmacher, M. Huber, L. Laff ers, H. Langen & M. Spindler, 'Causal Mediation Analysis with Double Machine Learning', *arXiv preprint arXiv:2002.12710*(2020). <https://arxiv.org/abs/2002.12710>.

32. Copus et al., above n. 22, at 50.

33. Dumas and Frankenreiter, above n. 28, at 63.

they are the experts in law. They are in the best position to understand the legal texts and to identify the interesting questions that can be answered using those texts. Therefore, CLA can benefit greatly from turning lawyers into users of computational methods. How to put it into practice is discussed in the next section.

The first article in this special issue – *Text-mining for Lawyers: How Machine Learning Techniques Can Advance our Understanding of Legal Discourse* by Arthur Dyevre – reviews in a non-technical manner some of the available computational methods (with the emphasis on the text-as-data approaches) and how they can be used in legal research. Even though this editorial also briefly reviewed some methods, this article enters into more detail about how those methods are applied and provides examples of the conclusions that can be derived from such research.

Computational methods require access to large amounts of data and the algorithms used to reach different conclusions often lack explainability and transparency (i.e., how the algorithm reached its outcome). This is of particular concern when computational methods are used not only in research but also in practice, for example, by decision makers in the criminal justice system and by financial institutions. Therefore, the second article in this special issue – *Big Data Ethics: A Life Cycle Perspective* by Simon Vydra, Andrei Poama, Sarah Giest, Alex Ingrams and Bram Klievink focuses on the ethical issues in using this data. In particular, it looks at the cycle of using big data, and which particular concerns are raised at each stage. For example, one of the common concerns with using predictive models to assess offenders' risk of recidivism is the inherent bias in past data (e.g., if one of the factors used by the algorithm for prediction of an increased risk is the ethnicity of the offender). This article is an important aid for remaining mindful of the concerns CLA brings with it, and for seeking to address those concerns.

The final article – *Teaching Technology to (Future) Lawyers* by Mi ok aj Barczentewicz addresses the 'elephant in the room': how are lawyers and legal scholars supposed to be able to apply these methods? The traditional law school curriculum, especially in Europe, does not provide for training in statistical or computational methods. This in principle limits the ability of legal scholars as well as practitioners to use these methods themselves. This article therefore discusses different models of training that can be introduced in order to enable such usage, at least to some extent.

These articles together attempt to present a more complete picture of CLA. Besides stressing the promises of these methods for legal research and practice, this special issue does not shy away from the challenges. However, stressing these challenges by no means suggests that promoting CLA among lawyers is a lost cause. It simply sheds light on the aspects that need to be addressed in order to better achieve such a goal. In the next section, I provide some thoughts on how, in my opinion, some of those challenges should be addressed and how CLA can be promoted among lawyers.

## 4 A Personal Note on the Way Forward

Here I would like to focus on the particular question of capacity: can lawyers fully utilise these tools given their limited expertise with such methods? Such discussion should differentiate between the current legal scholars, who have not been trained in these methods, and the future generations of lawyers and legal scholars.

Given lawyers' specialised knowledge of law and legal institutions, they can greatly benefit from, and also contribute to, the development of the field of CLA. They are in the best position to identify the relevant questions, provide the initial annotations and coding for training the models later, etc. However, most of the current generation of legal scholars, and especially the more senior ones, are not well equipped not only to use computational methods themselves, but also to understand them in a way that will allow them to come up with good research designs. Despite the small but growing group of legal scholars who are being trained in CLA, it cannot be expected that all legal scholars will become experts in computational methods, just as it was not reasonable to assume with the traditional empirical methods. Not only are the initial investment costs very high, but computational methods are also rapidly evolving and require constant expansion of expertise. It seems more reasonable to follow the logic of comparative advantages and solve the capacity 'problem' through collaboration rather than trying to capture everything. This new development (CLA) should be viewed as a great opportunity for legal scholars and methodology experts (who can come from different fields such as social sciences, digital humanities and computer sciences) to join forces. Therefore, the way forward for the current group of legal scholars is to facilitate such collaborative projects.

How can this be done? Even though one's own expertise in computational methods is not required for legal scholars in such collaborative projects, a basic understanding of the tools is necessary. As has been discussed in this editorial piece, there is a plethora of ideas that can be explored using these tools. However, in order to identify the research questions that are suitable for these techniques, the lawyers need to understand the possibilities and the limitations of the techniques. Law faculties could offer training programmes to lawyers to just understand the logic and the intuition behind each of the available methods. The participants in such programmes will not be required to develop any programming skills themselves or understand statistical models. The focus will be on a non-technical training, in which different examples can be used to demonstrate the nature of these methods. Once legal scholars understand what these methods are about, they can come up with their ideas derived from their deep understanding of the substantive legal fields. At this point, legal scholars can start collaborating with the methodology experts. The

training programmes will enable legal scholars to communicate with, for example, computer scientists using the proper jargon, thus allowing an easy and well-informed discussion. The methodology experts can then help by developing the necessary models and tools to execute the research. They can also comment on the potential limitations and add to the design of the research.

The expansion of the domain experts (legal scholars and lawyers) involved in this context can advance not only the research itself, but also the development of tools especially adjusted to the legal field. For example, methodology experts and engineers can develop NLP tools adjusted to law in different languages (given that national systems around the world have their unique legal languages). Furthermore, pulling together the routine efforts of lawyers while utilising computation power can save tremendous time and overlapping efforts. For example, annotating and coding case law in the course of legal analysis is a routine labour-intensive task made by many law students, lawyers and legal scholars. If such efforts are pooled together, machine learning models can be trained to make such annotations at scale. This will avoid duplication of effort; it will allow building a golden standard of annotation, which will serve not only research but also education. Better annotation software can then be developed, allowing for further annotation of similar legal documents. Finally, such an exercise can also benefit scholars from other fields who treat legal text as an additional object of research but lack the domain expertise to complete all the work themselves.

The proposal to develop such a training programme does not suggest that there will not be existing legal scholars who will choose to obtain that expertise themselves even in their advanced stage of legal career. Such legal scholars already exist, and this is a most welcome practice since they can enjoy both worlds. But if the entire field is to be built only on this small group, many opportunities for further developments will be missed. Therefore, I would suggest that investment should also be made to promote the understanding of the methods among the larger group of legal scholars who might resist (for obvious reasons) full scale (re)training.

The second group on which I would like to comment is the future generations of legal scholars/lawyers, who enter the law schools. An interesting discussion of, and suggestions for, the different opportunities and pitfalls of combining legal and technological education are put forward by Mikołaj Barczentewicz in the last article of this special issue. Therefore, here I only offer some general thoughts.

The increasing importance of technological literacy in general, and the future promise of CLA in particular, should render reforms in the law schools' curricula as an important and necessary step. However, the change needs to be made in such a way that complements the traditional methods of analysis rather than trying to replace them. Such an approach will take into account not only the interests of the students (it is doubtful that all incoming legal students will be interested in also

acquiring quantitative skills), but also of the labour market. It is not a new insight that law schools' curricula are heavily driven by market demands.<sup>34</sup> Even though there are already advances in legal services that utilise big data and information technologies, currently the legal labour market needs people who are trained in analysing and applying legal rules and cases. These skills are adequately provided by doctrinal education in law schools.<sup>35</sup> However, given the capacity of computational methods to complement legal analysis and make legal as well as scholarly work more efficient, soon enough acquiring such skills will provide an advantage to the graduating students. Already nowadays, top law firms employ lawyers with specific technology expertise. Moreover, in the future, technological literacy might even become indispensable for the legal field, but this is yet to be seen. An increased usage of legal text as data might also promote more extensive digitisation of legal documents by the respective authorities.<sup>36</sup>

In order not to lag behind, law schools can introduce parallel tracks (as some law schools already do, see Barczentewicz on this issue).<sup>37</sup> Alongside the standard law programme, an honours programme can be offered. In the latter, the students will receive in addition to the standard training in law, training in empirical methods (e.g., statistics, econometrics) as well as in computational methods (e.g., programming). Students can then choose for themselves whether to follow the technical training, thus, assuring self-selection of motivated and capable students. In the next step, a research master's in law with a focus on empirical legal studies in general and CLA in particular can be offered. Such a programme will build the human capital necessary to further develop the field. Furthermore, it will build in-house capacity in law schools, which will enable development and promotion of educational programmes without using external methodology experts. In the future, ELS in general and CLA in particular will be able to take a more prominent place in legal education if

the legal labour market (practice and academia) can utilise the new incoming human capital.

6

34. A. Dyevre, 'Fixing Europe's Law Schools', 25(1) *European Review of Private Law* 151-168 (2017); D. Hazel, M. Partington & S. Wheeler, *Law in the Real World: Improving Our Understanding of How Law Really Works, Final Report and Recommendations*(2006), at 29; R. Cooter, 'Maturing into Normal Science: The Effect of Empirical Legal Studies on Law and Economics', 2011(5) *University of Illinois Law Review* 1475-84.

35. Hazel et al., above n. 34, at 29.

36. For instance, currently, Dutch courts publish only a small portion of their decisions. An increased number of people applying CLA might create a demand-driven supply for legal texts. This can increase transparency and allow for important research to be conducted. Of course, a risk exists that precisely this will lead to the opposite reaction. In France, for example, the Government banned the publication of statistical information about judges' decisions. See [www.artificiallawyer.com/2019/06/04/france-bans-judge-analytics-5-years-in-prison-for-rule-breakers/](http://www.artificiallawyer.com/2019/06/04/france-bans-judge-analytics-5-years-in-prison-for-rule-breakers/) (last visited 8 August 2021). This decision followed the discontent of judges from NLP and machine learning companies who used public data to analyse patterns of specific judges' decisions.

37. For additional reading on different ideas how quantitative methodology can be introduced in legal educations, see also D. M. Katz, 'The MIT School of Law? A Perspective on Legal Education in the 21st Century', 2014(5) *University of Illinois Law Review* 101-42; A. Dyevre, *The Future of Legal Theory and The Law School of the Future*. Antwerpen: Intersentia (2015).