Introduction
Law as Data, Data as Law

Rónán Kennedy
University of Galway, Ireland
Brian M. Barry
Trinity College Dublin, Ireland

Abstract

The datafication of law is requiring a reconsideration of approaches to legal research and education. The academy is in an uncomfortable place, with new disciplines claiming to have analytical tools to offer and to be able to replace some or all of the traditional work of lawyers and teachers. This symposium issue provides an opportunity for reflection and contains contributions that range from a call to action which embraces the new technologies, critical perspectives on the proper place of data in legal research or in governmental practices, and empirical explorations of the effectiveness of large language models. It highlights themes of diversity, rigour and engagement as key issues for lawtech scholars to bear in mind in their future work and recommends that we embrace the discomfort so that we can mature with the field.

Law as Data, Data as Law: Drawing Breath Before Diving In

The theme of this symposium is *Law as Data, Data as Law*. It provides analysis from multiple disciplinary perspectives from commentators around the globe on the latest developments in data-driven approaches to law, and its impacts on legal practice, education and systems.

The motivation for this symposium issue is to facilitate measured and continuing reflection on the increasing datafication of law at a time of considerable change for law and legal systems. In particular, there has been rapid change in the capacity of information and communications technology to supplement the skills of lawyers (lawtech) – perhaps in part fuelled by the 'forced experiment' of increased digitisation of legal processes precipitated by public health restrictions during the COVID-19 pandemic.[1] Artificial intelligence (AI), particularly generative AI, has considerable implications for legal practice, the work of courts, and legal education. Law is increasingly subjected to quantitative analysis, while legal rules themselves are being 'translated' into software systems.[2] This may bring the practice of law from an individualised set of hunches and tacit intuitions to a more data-driven, scientific and objective approach, while creating the risk of fossilising it into opaque and difficult to challenge infrastructures.

Against this backdrop, we wanted to provide an opportunity to draw a breath. As the initial hype and hyperbole regarding generative AI fades somewhat, and scholars begin to highlight the problems and challenges of these tools,[3] particularly their negative social and environmental consequences, there is a need for lawyers to consider more deeply what the increased datafication of the discipline might mean for research agendas, teaching approaches, and the practice of law.

---

[1] Moran, "Business as [Un] usual."
[2] Frankenreiter, "Computational Methods in Legal Analysis"; Chau, "Computational Legal Studies Comes of Age."
[3] Mikalef, "Thinking Responsibly About Responsible AI."

Contributors to this symposium issue explore the overarching theme of *Law as Data, Data as Law* in a variety of ways. Certain contributions present new theories and frameworks for understanding these developments more holistically. Other contributions deploy emerging empirical methods to measure impacts on legal practice, education and systems.

Combined, these contributions build on the scholarly momentum investigating changes that emerged during and in the immediate wake of the COVID-19 pandemic. They showcase, in particular, the importance of methodological rigour and cross-disciplinary work to coherently conceptualise and evaluate in data-driven approaches to law.

The Contributions

Burgess et al. examine whether large language models are effective and in particular accurate in identifying a core skill of lawyers, that of identifying arguments and chains of argument in the reasoning of judges. When scored by a human expert using different methods, the various software tools had very variable performance, some barely passable but others with very high accuracy. They conclude that while these tools hold promise for the future, particularly in enhancing access to justice, law students cannot (yet) rely on them and human skills remain important.

Niamh Kinchin considers the detail of a particular application of *Law as Data, Data as Law* by unpacking the use of automated decision-making in refugee status determination. She lays out the use of computational legal reasoning – data as law in action – and closely explores the application of predictive analytics in the asylum process, highlighting a number of possible mismatches: between human reasoning and computational correlation; the abductive reasoning of status determination and the inductive reasoning of machine learning; the subjective perceptions of a fear of persecution in the present by a particular individual application and the collective experiences of past applicants reflected in training data. These issues and the challenges of 'data wrangling' mean that the choice of 'case characteristics' can have important implications for the correct application of the law – law as data in action – and that we must therefore interrogate algorithmic systems to mitigate against harmful bias and feedback loops. She concludes by calling for interdisciplinary dialogue in order to query these experimental technologies – 'shiny new things' that could bring negative transformations.

Audrey Plan moves the frame of reference for this symposium issue away from the common perspective on 'law as data' as being solely a matter of quantitative information and computational thinking. Instead, foregrounding the question of research methods, she argues that law has unique characteristics (non-commensurable texts, fluid concepts, unsystematic sources, and the resulting bias towards questions that can be answered rather than those which should) which require thoughtful approaches in order to address or mitigate these issues. This is a task which she encourages legal academics to embrace so that an appropriate research environment – indeed a methodological pipeline – can be developed.

Prakash and Nair put forward an urgent manifesto for legal educators to move beyond the Socratic method and integrate AI tools into the legal curriculum. This reform, the authors argue, should be grounded on ethical use, practical application, and interdisciplinary collaboration. Citing a range of studies that indicate that AI tools can enhance research efficiency, improve drafting precision, and create personalised learning experiences, the article calls for all involved to embrace a transformation with profound social impacts which can improve access to justice. It concludes with guidelines for future policy interventions and research to better support future lawyers to thrive.

Guillaume Zambrano explores the practicalities of working with law as data by testing the effectiveness and efficiency of different prompting strategies for large language models in extracting relevant cases from a dataset of court decisions. Even without optimisation, these achieve high performance metrics, and after optimisation, some improvements, although these are quite variable. This finding indicates that there is further work to be done on this topic and he concludes with pointers for future research questions.

Themes: Diversity, Rigour, Engagement – and Discomfort

Four themes emerge from a deeper consideration of these contributions. The first is diversity: not simply the diversity of topics considered, or approaches applied, but the new diversity that legal academics will increasingly be required to bring to their scholarship and their teaching. The call for papers elicited very divergent articles, ranging from a high-level call for greater integration of AI into the law school curriculum to a detailed critique of the application of AI in a specific context to hands-on quantitative testing of large language models for particular legal tasks. These bring not just new requirements of technical understanding and new research questions, but also new research methods. Reading Burgess et al. and Zambrano's papers in light of Plan's re-contextualisation of what we must do provides a more thorough understanding of the latter: asking only quantitative questions is *insufficient* but it is *necessary*. The first two articles show *how* this can be done; the challenge for the

legal academy is to acquire these new competencies without losing sight of what is different about law and to ensure that present and future scholars have a comprehensive and critical grounding in a methodological toolbox that is continuing to expand beyond doctrinal approaches, into socio-legal studies, and now into computational law.[4]

The second theme is the need for rigour – *why* both broad and detailed critique must be engaged in. Criticisms of socio-legal studies have a long history,[5] and although cross-disciplinary work can be stimulating and fruitful, it is not always grounded on sufficient expertise in multiple fields to be a true contribution to knowledge. It is again useful to juxtapose Burgess et al. and Zambrano with another contribution, this time that of Prakash and Nair. Rapid technological change is certainly creating new opportunities and inflection points that could bring about a better future for law graduates and for society as a whole, but there is a great deal of detailed work that legal researchers must engage in to ensure that the techno-utopian visions which are sold by vendors are subject to close examination and critique. In order to arrive at the better world that Prakash and Nair advocate for, studies will need to provide clear benchmarks that can be (re-)used to verify whether or not a particular tool is truly ready for real-world use. Kinchin's article is also relevant here: in line with Plan's perspective that the law should be the primary discipline that evaluates 'law as data, data as law' projects, but incorporating insights from other fields, Kinchin points to what could go wrong. She underlines the need to ensure that lawyers understand that they need to ask hard questions of technology that promises to make legal work quicker, simpler, or better.

The third theme is engagement – *who* this work is done for. Many contributors either hint at, or explicitly call for deeper engagement with both those who develop and deploy data-driven approaches to practising and teaching law, or hint at engaging with those impacted by those approaches. For example, Burgess et al highlight the perils of students' uncritical use of generative AI and the importance of educators' guidance to use these tools appropriately. The case for deeper scholarly engagement with educators and students on these issues would seem self-evident. Kinchin's nuanced treatment of the challenges of using automated decision-making tools to assess, among other factors, an asylum-seeking claimant's subjective fear of being persecuted underscores the critical role of human involvement (particularly legal experts' involvement) in algorithmic design. This resonates with broader arguments made elsewhere, proposing design thinking in law and legal systems – in essence, to listen to users' needs and experiences, and to embed their insights into the design of data-driven technologies for law.[6] Scholars have an important role to play in this regard: both to advocate and facilitate meaningful stakeholder- and citizen- engagement to appropriately develop and integrate more datafied ways of doing law.

The fourth theme is discomfort, which is *where* we find ourselves. Calls for cross-disciplinary dialogue are common in the law and technology literature, and it is not a surprise to see them appear in the contributions to this symposium issue. In order to elevate exhortations to look beyond our silos beyond a rote expectation, considering this group of articles side by side provides a more expansive and sophisticated understanding of what this will mean in practice for the legal academy. Editing this issue was not always a comfortable experience: authors draw on a wide range of different literatures. We were also taken out of our comfort zone by papers that were quite technical and required an evaluation that we are not accustomed to. It reminded us that, as a scholarly community, we cannot be complacent in our self-perception to be experts in the lawtech field. Indeed, it may indicate that the field has become so complex and so imbricated with other disciplines that true expertise is no longer possible (if it ever was). Computer science and information systems are two obvious examples of such disciplines, but in line with our focus on diversity above, we suggest that our readers consider what others should or could be added to that list. As the scope and scale of research on law and technology continues to expand, we must be prepared to be challenged not just by conclusions but by methods and arguments. In order for lawtech scholarship to mature, and to follow through on the requirements for diversity and rigour outlined above, we must all get comfortable with discomfort.

---

[4] Whalen, "The Emergence of Computational Legal Studies."
[5] Epstein, "The Rules of Inference."
[6] Ursel, "Building Better Law"; Toohey, "Meeting the Access to Civil Justice Challenge."

Bibliography

Chau, Bao Kham, and Michael A. Livermore. "Computational Legal Studies Comes of Age." *European Journal of Empirical Legal Studies* 1, no 1 (2024): 89-104. https://doi.org/10.62355/ejels.19684.

Epstein, Lee, and Gary King. "The Rules of Inference." *The University of Chicago Law Review* 69, no 1 (2002): 1-133.

Frankenreiter, Jens, and Michael A. Livermore. "Computational Methods in Legal Analysis." *Annual Review of Law and Social Science* 16, no 1 (2020): 39-57. https://doi.org/10.1146/annurev-lawsocsci-052720-121843.

Mikalef, Patrick, Kieran Conboy, Jenny Eriksson Lundström, and Aleš Popovič. "Thinking Responsibly About Responsible AI and 'the Dark Side' of AI." *European Journal of Information Systems* 31, no 3 (2022): 257–268. https://doi.org/10.1080/0960085X.2022.2026621.

Moran, Lyle. "Business as [un] Usual." *ABA Journal* 106, no 4 (2020): 34-41.

Toohey, Lisa, Monique Moore, Katelane Dart, and Dan Toohey. "Meeting the access to Civil Justice Challenge: Digital Inclusion, Algorithmic Justice, and Human-Centred Design." *Macquarie Law Journal* 19 (2019): 133-156.

Whalen, Ryan. "The Emergence of Computational legal studies: An Introduction." In "*Computational Legal Studies: The Promise and Challenge of Data-Driven Research*," edited by Ryan Whalen, pp. 1-8. Edward Elgar Publishing, 2020.

Ursel, Susan. "Building Better Law: How Design Thinking Can Help Us Be Better Lawyers, Meet New Challenges, and Create the Future of Law." *Windsor Yearbook of Access to Justice* 34, no 1 (2017): 28-59. https://doi.org/10.22329/wyaj.v34i1.4999.