

Introduction: Artificial Intelligence, Technology, and the Law

Simon Stern, University of Toronto Faculty of Law

simon.stern@utoronto.ca

[This is the pre-edit draft of an article that will appear in 68 University of Toronto Law Journal __ (2018); to cite or quote, please consult the published version]

On March 25, 2017, the Centre for Innovation Law & Policy at the University of Toronto hosted a conference on Artificial Intelligence, Technology, and the Law. The conference was supported by generous funding from the University of Toronto Law Journal and from the Social Science and Humanities Research Council of Canada. The contributions to this special issue of the UTLJ are based on papers originally presented at the conference. Some of the speakers discussed the kinds of tasks that Machine Learning (ML) and Natural Language Processing (NLP) can perform, when used to conduct legal research, to identify biases and discrepancies at the doctrinal level and in the performance of lawyers and judges, and to facilitate access to justice for those who cannot readily afford legal services. Other speakers considered the challenges that algorithms based on ML and NLP pose to democratic conceptions of legal authority. Taken together, the papers offered a range of views on the prospects and perils of AI for the practice of law and for the legal system as a whole. This introduction briefly describes the contributions, moving roughly from the more theoretical to the more concrete aspects of these issues.

Mireille Hildebrandt addresses fundamental issues concerning artificial intelligence and the conceptual foundations of law in her article, ‘Law *as* Computation in the Era of Artificial Intelligence: Speaking Law to the Power of Statistics.’ She observes that when commentators

discuss AI's disruptive potential for law, they tend to adopt an economic perspective, treating "the law" as equivalent with "the legal services market" and focusing primarily (often exclusively) on disruption to that market.¹ By contrast, she suggests that if one wishes to take an economic perspective, it is preferable to offer 'an analysis of the political economy that drives the "legal services market,"' and therefore to ask 'how law does and should constitute and regulate such economic markets, while sustaining the checks and balances of the Rule of Law that should define them.'² Instead of regarding law as a product or service, Hildebrandt explains, we might regard it 'as constitutive of social defaults' ensuring that 'those subject to law [have] effective legal remedies to contest a violation of their fundamental rights.'³ The operation of this system does not depend solely on enforcement (otherwise 'it would not be law but administration, discipline, or simply a matter of violence'); rather, it depends on that monopoly of violence as the means of underwriting 'the ... performative speech acts' whose articulation and enactment expresses and reaffirms the norms that bind the system.⁴ By understanding law as 'a coherent web of speech act that inform the consequences of our actions,' we can see the difference between 'the mathematical simulation of legal judgment' and 'legal judgment itself.'⁵

Elaborating the implications of this difference, Hildebrandt first observes that ML's opaque algorithmic procedures (whether the result of 'deliberate hiding of the source code' or concern to protect a firm's trade secrets and intellectual property rights), make it 'difficult or even impossible to test [the algorithm's] reliability and to investigate the assumptions that inform

¹ Mireille Hildebrandt, 'Law *as* Computation in the Era of Artificial Intelligence: Speaking Law to the Power of Statistics' (2018) 68: *UTLJ* __, at __.

² *Ibid* at __.

³ *Ibid* at __.

⁴ *Ibid* at __.

⁵ *Ibid* at __.

[its] development.’⁶ Moreover, ‘even if experts manage to verify the software, most of us lack the skills to make sense of it,’ rendering us dependent on its creators or credentializers.⁷ Second, she notes that statistical determinations, regardless of their accuracy, depend on an entirely different kind of ‘reasoning’ from the modes of analysis and justification essential to practices of legal argumentation: the former involves ‘numerical rearticulation and mathematical functions that connect the dots,’ while the latter involves justifications that present themselves as open to challenge.⁸ Besides pointing out this gap, Hildebrandt emphasizes that lawyers and those subject to law ‘should learn to speak the language of variables, functions, correlations, training sets, [and] hypotheses space,’ because only by doing so can we begin to discern the varieties of bias and distortion embedded in algorithms, thereby gaining the ability to ‘*speak law to the power of statistics.*’⁹ Third, the very means by which highly accurate, detailed, and specific legal algorithms are created may entail violations of various rights, including privacy rights and the right to non-discrimination.¹⁰ Last, the adoption of sophisticated legal machinery would likely result in the ‘gradual deskilling’ of those professionals who once performed these tasks, intensifying the problems noted above: just as authority is ceded to the algorithm, the experts best equipped to check its results find it increasingly hard to tell when a result ‘incorrect, imprecise or debatable.’¹¹

One question that Hildebrandt’s analysis prompts involves those places in the legal system that have typically relied on the scant mode of justification that she associates with the output of legal algorithms. Some kinds of administrative decisions are delivered in such a

⁶ Ibid at ___.

⁷ Ibid at ___.

⁸ Ibid at ___.

⁹ Ibid at ___.

¹⁰ Ibid at ___.

¹¹ Ibid at ___.

perfunctory manner as to appear nearly indistinguishable from such outputs. In most instances, they remain at least theoretically open to challenge, although their bland and inexpressive format may also serve, in practice, to discourage many people from inquiring into their basis. That these decisions are the products of a system that abides by Hildebrandt's requirements for reasoned justification does not necessarily allow them to avoid the criticisms she raises; hence her analysis might also suggest ways of comparing the features of an algorithmic system to those in a system that differs from it in principle, while coinciding with it in certain aspects.

Brian Sheppard also discusses theoretical questions involving the rule of law in his contribution, 'Warming up to Inscrutability: How Technology Could Challenge Our Concept of Law.' Noting that 'law is, like technology, a human artifact' that we adapt in response to changing needs, he asks whether we might, in time, come to tolerate certain kinds of 'systemic weaknesses'¹² arising from the routine use of technology—weaknesses that could, in turn lead us to revise our concept of law, fundamentally changing our conception of law's normative structure or the normative relation that defines our willingness to subject ourselves to legal authority. Could technology, Sheppard asks, 'change ... our perceptions regarding the delegation of official tasks to machines[?]'¹³ Focusing specifically on technology that could change our attitudes concerning the fashion in which we are regulated by law-like systems, Sheppard discusses the potential of ML to solve legal problems, and to produce legal directives, by way of 'automated individualized rule-making'—that is, by issuing rules specifically directed at a given recipient.¹⁴ Because ML proceeds by pattern recognition, to reach conclusions in a probabilistic fashion, these rules 'would not likely form an underlying structure of meaning with each

¹² Brian Sheppard, 'Warming up to Inscrutability: How Technology Could Challenge Our Concept of Law' (2018) 68: _ UTLJ __, at __.

¹³ Ibid at __.

¹⁴ Ibid at __.

other’;¹⁵ rather, they would form, in the aggregate, a mosaic comprised of numerous pieces, forming no coherent pattern. Nor would this patternless assortment even be accompanied by reasons that would be comprehensible to the recipients: whereas we prefer that legal decisions be justified ‘in intelligible language, reasonably short, and sufficiently comprehensive,’ the algorithmic outputs of the legal machine would likely be ‘impossible to interpret ... in human terms.’¹⁶ Central among the weaknesses of this system, then, would be its inability to tell us why we are required to follow a certain direction, why we should be held civilly or criminally liable for a certain course of conduct. We would have to trust in the system, effectively abandoning one of the central features of our legal system—namely, our ability to challenge a law’s validity.

After showing why a legal regime built on ML would likely have this weakness, Sheppard considers the problem of a legal system that is incoherent (to human perception, at least) from the perspectives of various theories of law. One way to see the dramatic shift that such a system would entail is to consider Hart’s theory, which requires that those whom the law governs must share ‘a critical reflective attitude to certain patterns of behaviour as a common standard,’ an attitude that ‘display[s] itself in criticism, including self-criticism, [and] demands for conformity,’ among other things.¹⁷ Could a machine adopt such a critical attitude? Probably not, Sheppard observes: ‘It will always speed along in service of [a specified] goal ... but it will not be able to step outside the system of primary rule generation to revise its overall goal,’¹⁸ and such revision is precisely what the critical stance would facilitate. Nor does a Razian view fare any better. For Raz, law’s claim to legitimacy depends on a claim to moral authority, which in turn depends on law’s ability to justify its requirements. Someone should comply with legal

¹⁵ Ibid at ___.

¹⁶ Ibid at ___.

¹⁷ Ibid at ___ (quoting HLA Hart, *The Concept of Law*, 3d ed (Oxford: Oxford University Press, 2012), 53).

¹⁸ Ibid at ___

directives if doing so ‘makes it more likely that the person ... will comply with right reason’ than if she acted on her own reasons. Since the legal machine would issue directives without reasons, and often without any discernible justification, someone attempting to ‘assess the legitimacy of the system’s claim to authority’ would often ‘have a difficult time knowing whether a conflicting reason is excluded or falls outside of its scope.’¹⁹ The system’s inscrutability fundamentally undermines its claim to authority, in Razian terms: ‘*Rational* action is not passive; it is not simply that the chosen action turned out to be the right one, but also that it was done for the right reason.’²⁰ Finally, a Dworkian approach would similarly demand that ‘the correct method of adjudication [must be] highly interpretive and evaluative,’ because, on Dworkin’s view, we must interpret the law ‘in its best light ... as if it were created by a community speaking with a single voice ... seeking to provide a justification for state coercion.’²¹ Not only does the legal machine offer no such justification, its ‘static nature’—that is, its inability to develop new norms, in the process of its operation—means that it cannot recognize and implement any new principles that ‘become evident through interpretation.’²²

If Sheppard indicates the significant gap between some of the prevailing theoretical accounts of law’s authority and the ground that would obtain in a system that depends on machine-generated legal directives, he also implicitly poses a question about how we might get from here to there, what it would take for us to ‘warm up’ to the state of affairs he describes. A system as comprehensive as the one he describes would not appear overnight; it would develop slowly and incrementally. Perhaps it would initially apply to legal domains that already appear opaque to most people, or that seem trivial and not in great need of careful justification. This is

¹⁹ Ibid at ___.

²⁰ Ibid at ___.

²¹ Ibid at ___.

²² Ibid at ___.

not to say that the opacity of certain complex and byzantine regulatory fields is the same as the inscrutability that Sheppard describes; however, because they are easily conflated, people might be more willing to trade the one seemingly incomprehensible routine for another, than if such an approach were to arise in an area such as constitutional law. Sheppard's analysis thus comes with an implicit warning, for those who consider that the legitimacy of law's authority depends on one or another of the theories he considers—or that, in any case, it depends on the justificatory requirement that forms the least common denominator of all three: namely, to monitor the early uses of a Machine Learning approach and its extension to other legal domains.

The contribution by Frank Pasquale and Glyn Campbell, 'Prediction, Persuasion, and the Jurisprudence of Behaviorism,' continues some of the themes of the first two articles, while also turning more concretely to some of the specific claims made on behalf of predictive algorithms. Pasquale and Campbell examine what they call the 'emerging jurisprudence of behaviorism,' which treats judicial decision-making according to a 'Skinnerian model of mental processes as a black-boxed transformation of inputs into outputs.'²³ Such an approach, they observe, jettisons the traditional legal concern with persuasion (based on explanation) and replaces it with prediction. As they point out, a commonly offered justification for the trade is that we are, in any 'strangers to ourselves' (as the psychologist Timothy Wilson puts it),²⁴ and hence the predictive approach merely exchanges one mode of inscrutability for another. But the two modes are different, as Pasquale and Campbell show: whereas a myriad of social and cognitive psychologists (among others) have been able to study the factors that influence judicial behavior, leading to 'more rigorous and open discussions of policy considerations motivating judgments,

²³ Frank Pasquale and Glyn Campbell, 'Prediction, Persuasion, and the Jurisprudence of Behaviorism' (2018) 68: _ UTLJ __, at __.

²⁴ Ibid at __.

and ... franker recognition of judges as political actors,'²⁵ the black-boxed nature of the algorithm's output would afford no such analysis. At the end of the day, judges can be 'questioned and rebuked for discriminatory behavior,'²⁶ whereas an algorithm subtly premised on biased data or premises could remain virtually immune from criticism. Turning to a specific treatment of the potential that NLP offers for this jurisprudence of behaviorism, Pasquale and Campbell consider the claims set out in a 2016 article by Nikolaos Aletras and his colleagues, which affords an NLP perspective on decisions in the European Court of Human Rights.

First and foremost, Pasquale and Campbell ask, what are the uses and purposes of such a study? The aims of Aletras and his colleagues are threefold: to take a first step in showing the value of NLP for making predictions about judicial decisions; to create a triage mechanism that could highlight those cases most appropriate for judicial response; and to show whether facts or law have a greater influence on a case's outcome. As to the study's predictive claims, Pasquale and Campbell question whether a sample that drew on the Court's own summation of the circumstances warranting judicial treatment should be credited with any 'predictive' power. The study did not look to the parties' own characterizations of the circumstances (prior to adjudication), nor to the circumstances as they might have appeared to third persons aiming to tell which fact patterns would be construed as amounting to human rights violations; rather the study drew on the descriptions offered by the Court, in the process of explaining its decision—a detail that is endogenous to the very outcome being predicted.²⁷ Additionally, the study excludes from its corpus those judgments that 'do not meet "admissibility" standards, which are largely procedural in nature,' and hence the 'cases where law entirely determined the outcome ... were

²⁵ Ibid at ___.

²⁶ Ibid at ___.

²⁷ Ibid at ___.

removed from a data set ostensibly representative of the universe of cases generally.’²⁸ Remove a group of cases that turn on issues of law, and the ‘facts’ will accordingly turn out to have more influence on the outcome, in the remaining set. Again, and in concert with some of the other contributors, Pasquale and Campbell note that statistical correlation, ‘as a pragmatic ethic of effectiveness,’ differs fundamentally from modes of analysis ‘associated with a rich history of meaning’;²⁹ thus one concern is that words that turn out to correlate well with a finding of a human rights violation, such as ‘son,’ may not be construed by the algorithm as an index of ‘family membership’ but would instead serve to ‘prioritize cases involving sons over daughters.’³⁰

In light of the various concerns about the way these data were compiled, and the shortcomings they exhibit for those who might wish to have mechanisms for prediction or triage, Pasquale and Campbell conclude that while statistical analysis can ‘shed light on troubling patterns of rulings’ and can help to reveal patterns of judicial bias, futurists should be more modest in their claims for the potential of systems based on NLP and ML.³¹ Even if predictive analytics were to become significantly more reliable, ‘the pragmatic and the critical uses of predictive algorithms are in tension,’³² precisely insofar as their pragmatic use may factor in biases of various kinds, whereas their critical use would treat that very possibility as a question for investigation.

Given that, as Pasquale and Campbell note, the efficacy of ML-based algorithms is generally at odds with the algorithms’ intelligibility for humans, a more general concern they

²⁸ Ibid at ___.

²⁹ Ibid at ___.

³⁰ Ibid at ___.

³¹ Ibid at ___.

³² Ibid at ___.

raise has to do with the ability of any algorithmic system to satisfy the needs of democratic legitimacy while also excelling at the legal tasks it performs. One implication of their analysis—consistent with Hildebrandt’s comments about ‘deskilling’—is that we may prefer to use ML as an aid rather than a substitute, not least because there are so many opportunities for inadvertent reliance on various forms of bias, in the design of algorithmic systems, that human oversight may remain a necessity.

Like Pasquale and Campbell, Paul Gowder discusses the manifestations of ML in concrete terms, in his article, ‘Transformative Legal Technology and the Rule of Law,’ but Gowder’s focus is on AI’s potential for traditionally underserved populations. Gowder considers two aspirations for legal technology—reducing the cost of obtaining legal information (‘Cheaper Lawyers’), and using ML, and other automated techniques, to change the playing field on which legal disputes are conducted (‘Transformative Artificial Legal Cognition’).³³ Whereas claims and predictions about the effects of ML usually involve various means of providing faster and less expensive legal services, Gowder focuses on the second option, considering strategies for using technology to address structural problems of unequal access to legal resources and to legal redress more generally. Kenneth E. Scott’s classic article ‘Two Models of the Civil Process’ proposed that class actions could be understood not only as collecting large numbers of individually litigable claims, but also as facilitating a different kind of lawsuit aimed at responding to harms that are cognizable only when viewed in the aggregate;³⁴ in the same spirit, Gowder suggests that automation can be used to mount legal challenges that would be inconceivable if formulated in individual terms. Noticing that particular types of ‘repeat

³³ Paul Gowder, ‘Transformative Legal Technology and the Rule of Law’ (2018) 68: _ UTLJ __, at __.

³⁴ Kenneth E. Scott, ‘Two Models of the Civil Process’ (1975) 27 Stan L Rev 937.

players³⁵ (e.g., landlords, banks) are typically the beneficiaries of proposals (and technologies already in use) for decreasing the cost of legal services through automation, Gowder draws on his experience when working in a low-income legal services office to posit an ‘egalitarian role’ for Transformative Artificial Legal Cognition.

Just as the class action facilitates the litigation of numerous aggregated low-value claims, not worth litigating individually, legal technology might allow consumers to respond to the boilerplate practice of setting out terms and conditions that amount to the ‘deletion of rights.’³⁶ That is, individuals might approve a feature providing that they will cancel their service, with their Internet Service Provider, unless the ISP ‘agree[s] not to sell consumer data to marketers.’³⁷ By ‘automat[ing] the performance of the threat,’ so that consumers agree to have it executed on their behalf only after some minimum number (e.g., one million) others also agree to do so, they might effectively ensure that the threat ‘need not be carried out, because the other party has an incentive to act in order to avoid the behavior to which the threat responds.’³⁸ Gowder imagines a doomsday device called ‘Dr. StrangeContract,’³⁹ by which consumers might opt for various conditions (regarding the sale of personal information, or mandatory arbitration, for instance) that would restore the unequal balance of power that boilerplate has been able to exploit. ‘Replicate that reasoning over a sufficiently large number of firms, and suddenly the balance of power in mass consumer contracts changes.’⁴⁰

³⁵ Although Gowder does not use this term, he develops a contrast that recalls the classic comparison between ‘repeat players’ and ‘one-shotters’ in Marc Galanter, ‘Why the “Haves” Come out ahead: Speculations on the Limits of Legal Change’ (1974) 9 Law & Society Rev 95.

³⁶ Gowder, *ibid* at ___.

³⁷ *Ibid* at ___.

³⁸ *Ibid* at ___.

³⁹ *Ibid* at ___.

⁴⁰ *Ibid* at ___.

A related approach might offer a new conception of ‘predictive policing.’ Whereas that idea is currently being deployed to identify high-crime areas, or to create profiles matching those who are likely to be engaged in criminal behavior, Gowder suggests that a police department’s ‘internal affairs divisio[n] ... could examine use of force records, as well as stop and arrest records of officers by race,’ which could then be used to ‘investigat[e] ... statistical outliers.’⁴¹ A police department taking a pre-emptive ‘public health’ approach might use ‘predictive modeling techniques’ to identify those officers who pose a significant risk of ‘having unlawful force or race discrimination complaints brought against them,’ giving them ‘additional training to head off the conduct before it occurs.’⁴² Nor, Gowder suggests, should these approaches be applied entirely at the individual level. Predictive analytics could establish that the ‘social norms of particular cultural groups within a jurisdiction’—norms relating to clothing and body language, for instance—are likely to be incorrectly identified as suspicious, and therefore to lead to ‘fruitless investigatory actions’ that would waste police resources and alienate those who are targeted.⁴³ Predicting such outcomes could, then, suggest the need for training to pre-empt it, and the ‘persistent failure to correct such errors’ might be treated as ‘evidence of system-wide bias.’⁴⁴ By elaborating these and other proposals, Gowder shows how technology might be used for various ends that are not usually contemplated by proponents of legal automation. He imagines a set of options whose systemic implementation would effect a very different kind of change from the one associated with reducing the cost of legal information.

Finally, concluding the issue is a contribution by Benjamin Alarie, Anthony Niblett, and Albert Yoon, whose work in developing a software tool that applies to ML to law informs their

⁴¹ Ibid at ___.

⁴² Ibid at ___.

⁴³ Ibid at ___.

⁴⁴ Ibid at ___.

article, ‘How Artificial Intelligence Will Affect the Practice of Law.’ They start by considering recent changes to the market for legal services and the prospects for automating certain tasks currently performed by lawyers. As they note, the practice of law is distinctive for the increasing complexity and specialization of the work it requires, the unpredictability of the time required to complete a task, and the client-oriented focus of a legal professional.⁴⁵ As clients have increasingly sought to economize on legal services during the last decade, legal practice has evolved in various ways to meet this demand, and the ‘technical progression of legal information’ has offered one means of responding.⁴⁶ Over time, the format in which legal information is stored and accessed has changed from analog, to digital, to computational, and this most recent turn has, of course, been the focus of the contributions to this issue. ML, in particular, offers an extremely powerful tool for legal research, given its ability to ‘develop tools that identify nuanced patterns in data, beyond what highly skilled, experienced workers could reasonably construct on their own.’⁴⁷

One of the specific applications of ML to law involves tools for e-discovery, which replaces a ‘highly labour-intensive’ activity requiring ‘a small army of lawyers’ with an automated process that can be ‘conducted with a fraction of the time [and] expense ... [and] with even greater accuracy.’⁴⁸ Other applications involve the use of NLP to create forms and contracts that ‘avoid[d] many of the pitfalls of drafting, such as ambiguous conditions and omission of key provisions or terms,’ and to identify new caselaw relevant to a client’s interests, in turn providing gist for a memo that explains the law and advises the client accordingly.⁴⁹ Blue J Legal, the

⁴⁵ Benjamin Alarie, Anthony Niblett, and Albert H. Yoon, ‘How Artificial Intelligence Will Affect the Practice of Law’ (2018) 68: _ UTLJ __, at __.

⁴⁶ Ibid at __.

⁴⁷ Ibid at __.

⁴⁸ Ibid at __.

⁴⁹ Ibid at __.

startup organized by the article’s three authors, focuses on tax law, and advises clients by ‘identify[ing] a fact-intensive question of law’ according to factors that bear on taxation of various tasks and activities, then ‘cod[ing] every published decision’ in terms of those factors, and finally ‘provid[ing] a written report’ that describes the likely legal classification, the degree of confidence for that answer, an explanation for the result, and the caselaw that supports it.⁵⁰ As the authors note, not all areas of law are equally amenable to this approach. Over time, however, the great availability of data may help to facilitate its spread to other legal domains. Just as such tools may help with the practice of law, they may also, in time, come to serve various judicial and legislative needs—allowing lawmakers, for instance, ‘to make laws that are better tailored to individual circumstances.’⁵¹ Although the authors do not predict that such a transformation is imminent, they suggest that with time, ML and related technologies ‘will improve transparency and empower lawyers to work more efficiently deepen and broaden their areas of expertise, and provide greater access to justice and more value to clients.’⁵²

It remains only for me to thank David Dyzenhaus and Karen Knop for their enthusiastic response to my initial ideas about the conference; Leah Theriault, the former Administrative Director of the Centre for Innovation Law & Policy, for her superb assistance in organizing the conference, and Ellie Marshall, my Research Assistant, for her valuable help with many aspects of the conference and this special issue. I would also like to express, once again, my gratitude to both the UTLJ and the Social Science and Humanities Research Council for their support.

⁵⁰ Ibid at ___.

⁵¹ Ibid at ___.

⁵² Ibid at ___.