

Argument by Numbers: the Normative Impact of Statistical Legal Tech

Laurence Diver¹
Pauline McBride²

ABSTRACT

The introduction of statistical ‘legal tech’ raises questions about the future of law and legal practice. While technologies have always mediated the concept, practice, and texture of law, a qualitative and quantitative shift is taking place: statistical legal tech is being integrated into mainstream legal practice, and particularly that of litigators. These applications – particularly in search and document generation – mediate how practicing lawyers interact with the legal system. By shaping how law is ‘done’, the applications ultimately come to shape what law *is*. Where such applications impact on the creative elements of the litigator’s practice, for example via automation bias, they affect their professional and ethical duty to respond appropriately to the unique circumstances of their client’s case – a duty that is central to the Rule of Law. The statistical mediation of legal resources by machine learning applications must therefore be introduced with great care, if we are to avoid the subtle, inadvertent, but ultimately fundamental undermining of the Rule of Law. In this contribution we describe the normative effects of legal tech application design, how they are potentially (in)compatible with law and the Rule of Law as normative orders, particularly with respect to legal texts which we frame as the proper source of ‘lossless law’, uncompressed by statistical framing. We conclude that reliance on the vigilance of individual lawyers is insufficient to guard against the potentially harmful effects of such systems, given their inscrutability, and suggest that the onus is on the providers of legal technologies to demonstrate the legitimacy of their systems according to the normative standards inherent in the legal system.

KEYWORDS: automation bias, legal technology, machine learning, creativity, Rule of Law

This work has received funding from the European Research Council (ECR) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 788734).

RÉSUMÉ

L’introduction des “*legal techs*” statistiques amène avec elle des questions sur le futur du droit et celui de la pratique juridique. Alors que depuis

toujours les technologies affectent le concept, la pratique et la texture du droit, nous pouvons constater désormais un changement qualitatif et quantitatif : les *legal techs* statistiques sont intégrées dans la pratique juridique courante, particulièrement en matière de litiges. Ces applications – en particulier celles qui automatisent la recherche juridique et la génération de documents – influencent la façon dont les avocat.e.s praticien.ne.s interagissent avec le système juridique. En définissant comment le droit se « fait », ces applications finissent par définir ce qu'est le droit. Ces dernières impactent la créativité des avocat.e.s en litige dans leur pratique, par exemple à travers l'automatisation des biais, et affectent par là-même leur devoir professionnel et éthique qui consiste à offrir une réponse appropriée aux circonstances particulières du cas de leur clientèle – c'est un devoir qui est central à l'État de Droit. La gestion statistique des ressources juridiques à travers des application d'apprentissage automatique doit donc être introduite avec une attention particulière afin d'éviter d'ébranler de manière certes subtile et involontaire, mais aussi fondamentale, l'État de droit. Dans cette contribution, nous décrivons les effets normatifs des applications en *legal tech*, ainsi que leur potentielle (in)compatibilité avec le droit et l'État de droit comme ordres normatifs, particulièrement au regard des textes juridiques que nous considérons être la source du droit préservé (*lossless law*), du droit qui n'est pas amputé par les approches statistiques. Nous concluons qu'afin d'éviter les potentiels effets néfastes des systèmes statistiques en droit, notamment liés à leur impénétrabilité, il n'est pas suffisant de compter sur la seule vigilance des avocat.e.s. Nous suggérons donc qu'il revient aux fournisseurs des technologies juridiques de prouver la légitimité de leurs systèmes, sur la base des standards normatifs inhérents au système juridique.

MOTS-CLÉS : automatisation des biais, technologies juridiques, apprentissage automatique, créativité, État de Droit

Ce travail a été financé par le Conseil Européen de la Recherche (ECR), sous l'égide du programme de recherche et d'innovation Horizon 2020 de l'Union européenne (No 788734).

¹ Postdoctoral researcher, Counting as a Human Being in the Era of Computational Law (COHUBICOL), Law, Science, Technology & Society research group (LSTS), Vrije Universiteit Brussel

² Postdoctoral researcher, Counting as a Human Being in the Era of Computational Law (COHUBICOL), Law, Science, Technology & Society research group (LSTS), Vrije Universiteit Brussel

Introduction

The introduction into legal practice of new technologies based on statistical prediction raises questions about the nature of legal normativity. Although it has always been the case that technologies have mediated the concept, practices, and commitments of law, there is a qualitative and quantitative shift taking place as data-driven “legal tech” is introduced into mainstream legal practice, and particularly that of litigators. The affordances of these applications necessarily mediate the interactions of practising lawyers with the legal system. By shaping how law is “done,” the choice architectures of the applications involved come to shape what law *is*. This implies the potential of legal tech to affect the normative structure of the law, and its capacity to balance a plurality of political *goals* while simultaneously restricting the arbitrary exercise of political *power*. Even though legal tech will in some cases facilitate more efficient and cost-effective legal practice,¹ there is a sense that, at least in some forms, it represents the subtle injection of a kind of political rationality directly into the heart of the legal process.² This matters, because in principle, legality and the Rule of Law ought to come *before* such technologized expressions of political morality, rather than be constituted by them.

The European Commission’s proposal for an AI Act hints at what is at stake. As presently drafted, the Act classifies as “high risk” those systems that are “intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts”; such systems are classified as being involved in the “Administration of justice and democratic processes.”³ The concern (and these provisions) ought to extend beyond judicial authorities, however. Given the welcome recognition in the AI Act of the novel challenge posed by computational systems to the administration of justice and to democracy, its ambit ought to apply more widely to encompass the uses of AI by parties whose professional role in the administration of justice obliges them to sustain a normative commitment to legality and the Rule of Law. A central element of the lawyer’s credo is the “will to justice,” an ethical and professional commitment to the governing ideal of law that must motivate day-to-day legal practice, even and especially where the practical realities of the latter make it difficult to sustain.

That commitment can be distorted when the mediating interface between lawyer and legal system is a software application whose design

¹ Few would argue seriously that we should go back to using typewriters instead of word processors, for example—although the same might not be said for email.

² Cf. Rouvroy and Stigler’s discussion of “algorithmic governmentality.” See A. Rouvroy and B. Stiegler, “The Digital Regime of Truth: From the Algorithmic Governmentality to a New Rule of Law” (2016) 3 *La Deleuziana: Online Journal of Philosophy* 6.

³ Proposal for an Artificial Intelligence Act (COM [2021] 206 final), Annex III, Article 8.

reflects not the will to justice, but a will to something else.⁴ The lawyer's role is to act as the agent for the client's interests, while at the same time maintaining fidelity to justice as an "officer of the court."⁵ The commitment this reflects is mediated by the tools they use in the course of their practice. Depending on how these tools are designed, the lawyer may find themselves deskilled, and their expertise and commitment as a lawyer undermined, if they are cast merely as a "user" of a legal technology whose design reflects an instrumentalized notion of law. The generative process of reasoning that is an inherent and creative element of their practice is potentially displaced, or even elided, in turn sidestepping the normative commitment to valid legal reasoning and processual legitimacy upon which the fundamental notion of "justice being seen to be done" is based.

For example, when a lawyer uses legal tech that delivers a first draft of, or otherwise structures or informs their written motions, pleadings or arguments, they allow the legal tech to shape those documents and arguments. Even if the legal tech does not have the final last word on the documents they prepare and submit to the court, they might let it set the starting point, pay attention to its suggestions, rely on its output to inform, if not direct, their strategy. Why is this problematic? If using the legal tech results in motions or pleadings which better capture the relevant law, or, in the case of arguments, contain more relevant citations, why should the use of the technology give cause for concern? There are two answers to this. The first is: How do we assess whether the results are "better"? We do not address that question here. The second concerns the role of lawyers, the nature of legal argumentation and the function of law itself.

In this contribution, we consider how the creative task of legal reasoning and attribution is mediated by legal technologies. We highlight the reflexivity of technologies used in legal practice, and how they can come to constitute the practices of which they are part. This leads us to draw attention to the role of the lawyer as, fundamentally, a creative marshal of hermeneutic and normative resources. As Gaakeer puts it, the role of the lawyer is to *fabulate*,⁶ to create stories that consist of not just the "data" of past cases, but also the expressions and intuitions of principle and the will to justice that (should) govern the whole enterprise. Their task is to metabolize these elements and their incomputable contributions into a new piece of legal reasoning that is inherently creative—even if only to the extent that it reflects a point-in-time abstraction or rationalization of the particular circumstances at which it is directed. The technological mediations of legal tech might, depending on their character, threaten this creative aspect of the litigation lawyer's role. The

⁴ Cf. I. Kerr, "Prediction, Pre-Emption, Presumption: The Path of Law After the Computational Turn" in M. Hildebrandt and K. De Vries (eds), *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology* (London: Routledge, 2013).

⁵ *Ibid.* at 13. See also A. Paterson, "Duties to the Court," *Law, Practice & Conduct for Solicitors* (2nd ed., W Green/Thomson Reuters, 2014).

⁶ J. Gaakeer, *Judging from Experience: Law, Praxis, Humanities* (Edinburgh University Press, 2019).

very possibility of a creative response is central to the professional and ethical duty the lawyer is under to respond sensitively to the unique constellation of facts, norms, and arguments that the enterprise of law charges them with asserting, defending, and countering.

Our claim is that the normative choice architectures presented by the design of legal tech applications must complement and uphold the implied philosophy of law. Legal professionals—not just judges—are not mere observers of a data-driven production line but are vocational practitioners. Their normative commitment to legal protection is manifested *inter alia* via the creative linguistic synthesis of the past and the present, always with an eye toward the future. When the “third voice” of a legal tech system that is predicated on *statistical* rather than *legal* relevance is inserted into this practice, it distorts or refracts the relationship between the “raw” text of the law and the creativity, experience, and commitment of the lawyer—potentially in ways that are not apparent, even to the lawyer themselves. The quality of that relationship is central to the concept of law, and to the efficacy of the Rule of Law. Any alterations to it must therefore be made in full awareness of the implications, especially where automation bias and the apparent effectiveness of new technologies are likely to be a deciding factor.

Normative Transformations: Technology’s Mediation of Law

“Legal tech” as a label has a relatively recent genesis, although there have been attempts to involve computational systems in the practice of law since the post-war period. It refers generally to the introduction of digital technologies into the practice, and occasionally the enforcement, of law. Legal tech systems have many disparate functionalities and aims, ranging from the very specific and verifiable (for example, the automation of contract checking on the basis of a firm’s playbook⁷) to the somewhat more difficult to substantiate (for example, “improving access to justice”⁸). Vagueness and commercially driven hype aside, it is important to appreciate that “technologies of law” are not a new phenomenon. As Hildebrandt notes, the phenomenon and practice of law have been facilitated for roughly the past five centuries by the technologies of the script and the printing press.⁹ Without those, law-as-we-know-it would be entirely unrecognizable from the practices and textual and material artifacts familiar to contemporary lawyers and citizens.

⁷ e.g. LawGeex <<https://www.lawgeex.com>> accessed 1 November 2021.

⁸ e.g. DoNotPay, whose terms and conditions state: “Our goal is to level the playing field and make legal help accessible to the most vulnerable in society” (see <<https://web.archive.org/web/20210929105038/https://donotpay.com/learn/terms-of-service-and-privacy-policy/>> accessed 1 November 2021).

⁹ Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar Publishing 2015) *passim*. See also D.J. Harvey, *Collisions in the Digital Paradigm: Law and Rule Making in the Digital Age* (Bloomsbury, 2017).

The technology of text provides constitutional democracies with a set of foundational *affordances* that are, reflexively, constitutive of such states. Text externalizes legal norms, giving them stability across time and space, in turn facilitating the notion of a jurisdiction that can be governed across time and space by that same set of texts. The inherent ambiguity of the natural language contained in those texts forces their interpretation at the point of application, confronting one citizen's particular view of the world with the content of the norm, resulting in a normative synthesis that is capable of guiding action. Therein lies the seed of contestation: The norm can be interpreted in different ways, thus affording—and perhaps sometimes provoking—competing narratives about its meaning, and how it ought to be applied in this case, in this community, at this particular moment in history. The potential for conflict that is inherent in the nature of text necessitates some form of adjudicative process that can decide on disputes of interpretation. Such decisions must in turn be legitimized within the community, which in constitutional democracies is achieved by means of the publicity of both the processes and the reasoning of judgments. By dint of these characteristics, legal judgments—which are themselves textual, and are thus fed back reflexively into the whole enterprise—serve as a marker of progress, and of the community's view of itself at a given moment, until the next phase of contest, adjudication, and judgment comes around.

The Importance of Creative Argumentation

The role of lawyers is to act as agents for the people they represent. Their primary duty is to their client, although they also owe duties to the courts and their fellow lawyers.¹⁰ Though lawyers draft motions, pleadings, and arguments, these are their *client's* motions, pleadings, and arguments. Like Waldron, we consider that the function of law is to enable such argumentation, to allow ordinary people, through their lawyers, to “fram[e] their own legal arguments, by inviting the tribunal hearing their case to consider how the position they are putting forward fits generally into a coherent conception of the spirit of the law.”¹¹

In Waldron's account, what matters is the *process* of argumentation. It is this which ensures “respect for the freedom and dignity of each person as an active intelligence.”¹² The decisions of courts, therefore, should be regarded not as mere “win or lose” outputs, but as the results of a process that involves listening to people's views, while constraining them to formulate those views against the backdrop of relevant legal norms.¹³ In the course of an

¹⁰ Paterson, *supra* note 5.

¹¹ J. Waldron, “The Rule of Law and the Importance of Procedure” (2011) 50 *Nomos* 3 at 19.

¹² *Ibid.* at 23.

¹³ The legal institutional process that facilitates such argumentation, Waldron points out, has a “dignitarian” aspect. Law treats people with respect. This involves “paying attention to a point of view and ... respecting the dignity of

adjudication, the litigator brings together the elements of a story that is compatible with the syllogistic framework of legal reasoning (if not necessarily presented as such).¹⁴ Appeals are made to certain norms, interpreted in light of the facts as presented, just as the facts are interpreted in light of the norms.

It is by no means easy to cleanly map states of affairs in the world (facts) to the requirements of legal norms. Evidence must be elicited and interpreted before being framed in linguistic terms that are intelligible more generally within the form of life of the law.¹⁵ Even setting aside the scientific and technical challenges of gathering evidence reliably, qualifying it in legal terms is not simply a case of description, even if using legal terminology, but is necessarily one framed in terms of the argument it is intended ultimately to support.¹⁶ Here rhetoric plays a role in setting out (i) exactly which facts are relevant, (ii) what aspects of them are relevant, in relation to (iii) which particular norms from the whole corpus of law, and (iv) why, in terms of the interests of the client and of justice. The rhetorical presentation of the argument will often hide its underlying syllogistic logic, but the necessarily creative aspects of the linguistic framing prevent the possibility of simple logic encapsulating all that is going on.¹⁷ There is no platonic notion of a logical “legal truth” that already exists “out there,” ready to be found either by the lawyer or by a machine. Instead, the creative work of interpretation and performative argumentation, undertaken by the litigator who is aware of their professional and ethical duties, is exactly what is entailed in “doing law.” It is only in making an argument, creatively synthesized from its constituent parts, that law’s values are sustained, and this is true independently of the outcome.

Two Problems of Legal Tech

In light of the brief characterization of creative legal argumentation just advanced, the problem of legal tech’s influence on the lawyer’s creative contribution can be considered along two distinct but interconnected dimensions. The first is the concern that computers cannot reason in the specific way that lawyers do, and that they do not and cannot have any normative commitment to the task at hand.

those to whom the norms are applied as *beings capable of explaining themselves*.” (*Ibid.* at 16, emphasis supplied). See also Cf. Y. Meneceur and C. Barbaro, “Artificial Intelligence and the Judicial Memory: The Great Misunderstanding” [2021] *AI and Ethics* at 3.

¹⁴ Cf. N. MacCormick, *Rhetoric and the Rule of Law: A Theory of Legal Reasoning* (Oxford; New York: Oxford University Press, 2005) ch. 3–4.

¹⁵ On this process of framing, see N. van Dijk, “The Life and Deaths of a Dispute: An Inquiry Into Matters of Law” in K. McGee (ed.), *Latour and the Passage of Law* (Edinburgh University Press, 2015); B. Latour, “Biography of an Inquiry: On a Book About Modes of Existence” (2013) 43 *Social Studies of Science* 287, ch. 13.

¹⁶ On the notion of the legal argument necessarily proceeding from within a legal frame, see “Working on the Chain Gang: Interpretation in the Law and in Literary Criticism” in S.E. Fish, *Doing What Comes Naturally: Change, Rhetoric, and the Practice of Theory in Literary and Legal Studies* (Durham: Duke University Press, 1989) at 100.

¹⁷ As MacCormick notes, it may not be “necessary or even best to set out real legal arguments in rigorously syllogistic form. How to present the case belongs more to rhetoric than to logic, but *the most effect rhetoric is likely to be that which rests on a clear understanding of the implicit logic of the process*,” *supra* note 14 at 42–43 (our emphasis).

The second dimension is more subtle, because its challenge to the status quo of legal practice is not so easily appreciated: It is the notion that the outputs of legal tech systems—in particular those built round machine learning prediction—will be treated as a definitive statement or reflection of the law itself, thereby affecting the underlying structure of law in ways that might be difficult to anticipate or even observe once those outputs are subsequently acted upon, and thus legitimized, within the legal process.

Computational Legal Reasoning

The statistical models at the heart of data-driven legal tech predict inductively as an exercise in pattern matching, while code-driven and symbolic systems follow deterministic rules to provide a solution whose essential structure was already anticipated in the design. Even where an output or prediction happens to be welcome, in neither case is there any necessary connection to normative legal practice, much less a *commitment* to those parts of it that require interpretative synthesis. In the case of statistical prediction, any good result is only good because we make the effort to interpret it as such, and our ability to do this is only because the result falls within a range that is not so totally outrageous as to make it impossible to interpret as reasonable. The machine has no access to standards against which we can test such an output for its reasonableness, since it has no notion of the value or normative context of the task it performs. In a sense, then, we are lucky when statistical models used in legal tech produce predictions to which we can impute a reasonable interpretation, since these models possess no normative stance guiding the prediction toward a just end. Simply put, when it comes to law, the statistical models “do not know what they are talking about.”¹⁸

Throughout the twentieth century, much energy was expended in the pursuit of *jurimetrics*, viewing law as a “science” and legal norms as logical premises that are susceptible to an austere form of deductive calculation. Folded in with the ethos of logical positivism and behaviourism (if not exactly their historical timelines), the future “man of statistics” was seen as the ideal jurist, running the legal “calculus” to derive the correct answer to any dispute.¹⁹ This view is clearly demonstrated in some contemporary legal technology applications and research agendas that seek to predict the outcomes of cases on the basis of extralegal data, for example the speech patterns of litigators.²⁰

¹⁸ B. Cantwell Smith, *The Promise of Artificial Intelligence: Reckoning and Judgment* (Cambridge, MA: The MIT Press, 2019) at 76 (emphasis supplied).

¹⁹ Gaakeer, *supra* note 6, ch. 3. The notion of the “man of statistics” in law comes from O.W. Holmes, “The Path of the Law” (1897) 10 *Harvard Law Review* 457.

²⁰ See the studies in M.A. Livermore and D.N. Rockmore (eds), *Law as Data: Computation, Text, and the Future of Legal Analysis* (Santa Fe: SFI Press, 2019). Note, however, that France has imposed legal limits on the use of data pertaining to judges’ behaviour. See F. G’sell, “Predicting Courts’ Decisions Is Lawful in France and Will Remain So” <<https://gsell.tech/en/predicting-courts-decisions-is-lawful-in-france-and-will-remain-so/>> accessed 30 September 2021.

These more contemporary data-driven approaches are built on the kinds of inductive inference enabled relatively recently by access to vast datasets, and by the computing power required to generate the statistical models that make the predictions possible. By contrast, since the 1970s, work in symbolic AI and law has sought to model and emulate the deductive reasoning capabilities of the lawyer. In the latter case, the aim is not to statistically predict the future on the basis of past data,²¹ but instead to predetermine the outcome through an *ex ante* specification of the “correct” answer contained in formalized rulesets that either present it or directly enforce it through additional code-based architecture. Here we see a crossover with another line of thinking that has arisen in the past 25 years, “code as law,” and the idea of a perfect and hyper-efficient enforcement of legal norms that promotes certainty as a core—perhaps *the* core—goal of the legal system.²² This instrumentalized view of law is enjoying a new lease of life in parts of the burgeoning “Rules as Code” movement.²³

Statistically Mediated Sources of Law

On the one hand, then, we have the problematic notion of the machine being in some sense an independent agent to which we “outsource” legal reasoning. On the other, we have a more subtle problem, one where the lawyer is not replaced *per se*, but instead relies on the output of the legal technology as if it were, or proceeded upon, an authoritative statement of the law, as in the case of legal search and motion drafting systems built around natural language processing.

The Lure of Effectiveness

One might think that lawyers are unlikely to be persuaded by the marketing efforts of legal tech firms, believing that their intuitions about the limits of computation and their understanding of the central role they play in the legal system will provide a source of healthy scepticism. Without presuming too much, however, it may be that commercial concerns succeed in incentivizing the adoption of certain technologies, especially where they purport to reduce cost and increase productivity. Resisting the lure of apparent effectiveness might be difficult for some, especially where network effects result in a majority of firms using the same or similar systems. As U.S. litigation analytics firm Lex Machina puts it, “Your competition uses Legal Analytics to

²¹ M. Hildebrandt, “Code-Driven Law: Freezing the Future and Scaling the Past” in S.F. Deakin and C. Markou (eds), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence* (Oxford, UK; New York, NY: Hart, 2020).

²² Z. Bańkowski, “Don’t Think About It: Legalism and Legality” in M.M. Karlsson, Ó. Páll Jónsson and E.M. Brynjarsdóttir (eds), *Rechtstheorie: Zeitschrift für Logik, Methodenlehre, Kybernetik und Soziologie des Rechts* (Berlin: Duncker & Humblot, 1993).

²³ Cf. M. Waddington, “Research Note. Rules as Code” (2020) 37 *Law in Context. A Socio-legal Journal* 1 (arguing for a restricted notion, and application, of Rules as Code).

win the client and the case; can you afford NOT to?"²⁴ Here we see a potential convergence of economic, legal, and technological cultures, where both the design and marketing of legal tech reflect a bias toward a commercially attractive but essentially utilitarian notion of law, a notion based chiefly on *outcomes* that are achieved through innovation and efficiency (and that, of course, generates healthy profits for the legal tech sector).²⁵

Beyond such marketing efforts, however, is the problem that, on the face of it, such systems do in fact work—at least in the sense that they generate intelligible outputs that can be acted upon. Accepting providers' claims of effectiveness may not seem like such a huge leap when it is evident that a system does in fact provide relevant search results, or generates wording for an argument or motion that is not wholly unreasonable. Statistical legal search systems like Casetext and Westlaw Edge quite clearly provide workable results; given a certain query, they will usually return results that the practitioner can make sense of, some of which will be as relevant as those that older systems or a manual search would have uncovered (and some even more so). Paradoxically, this is precisely where the problem lies. The mere existence of a workable output is attractive, particularly if it has the aura of "betterness" about it. The apparent success of the statistical output—bolstered by its speed and by claims made by its provider about the huge numbers and kinds of variables that are factored into its design—creates the sense that it has access to a better version of law—one that transcends the limits of human capabilities. Given the professional duty to represent the client's interests as best they can, the conscientious lawyer might even come to believe they *must* use such an apparently effective system—especially if professional guidance is silent on the broader and deeper implications of adopting such software.²⁶

Automation bias can have a beguiling effect: The system seems to work, it uses "objective" data (and much more of it than a human could possibly make sense of), and so there must be something of value in the outputs it generates that goes beyond the comparatively limited capacities of the human being.²⁷ The lawyer who goes against such apparent capabilities must

²⁴ "THE FUTURE OF LAW Judge Analytics: Predicting the Behavior of the Courtroom's Boss" (*Lex Machina*) <https://pages.lexmachina.com/Website_-Future-of-Law-Webcast_The-Future-of-Law-3.html> accessed 31 October 2021. The same firm claims that "more than three quarters of Am Law 100" (the top 100 firms in the United States, identified by American Lawyer magazine) use their system. See "Lex Machina: Legal Analytics® for the Data-Driven Lawyer" (*Lex Machina*) <<https://lexmachina.com/wp-content/uploads/Legal-Analytics-for-Data-Driven-Lawyer.pdf>> accessed 31 October 2021.

²⁵ Cf. D.M. Katz, "Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry" (2012) 62 *Emory Law Journal* 909. On the conflation of "law" with the "legal services market," often associated with the University of Chicago and the Law and Economics school that it birthed, see M. Hildebrandt, "Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics" (2018) 68 *University of Toronto Law Journal* 12 at 12–15.

²⁶ For example, the Law Society of England and Wales refers to the "rule of law" only in terms of a legal tech system's compliance with "all applicable laws," with no mention of the potential for deeper effects on practice and the concept of the Rule of Law. See "Lawtech and Ethics Principle" (Law Society of England and Wales, 2021) at 13–14 <<https://www.lawsociety.org.uk/topics/research/lawtech-and-ethics-principles-report-2021>> accessed 30 October 2021.

²⁷ On the problems for the law of relying on artificial intelligence outputs, both in terms of liability and of evidence, see S. Mason, "Artificial Intelligence: Oh Really? And Why Judges and Lawyers Are Central to the Way We Live Now—but They Don't Know It" [2017] *Artificial Intelligence* 29.

be very sure of their own position; the kind of expertise that until now would underpin such certainty might even come to be viewed as “obstinacy, arrogance, or presumption.”²⁸ In the long-term, we might start to see in the algorithm an idealized version of what we could be, if only we were able to retain as much information, correlate as many features, and perform processing at such immense speed. This “robotomorphy” is the opposite of anthropomorphism: We start to see the machine in ourselves and, owing to its apparent excellence in performing its task, we try to “optimize” ourselves and our practices in an attempt to align with that algorithmic ideal.²⁹

This idea is reflected in behaviourist conceptions of law that elide any internal normative commitment to what is necessarily a human practice. Volokh, for example, suggests that a hypothetical robot judge that could write opinions that are intelligible and persuasive ought to be deemed eligible to perform real judgments in real legal proceedings.³⁰ There, the test is one of building a system that succeeds in producing a judicial simulacrum, rather than one that can stand in the judge’s intentional shoes, guided by the normative will to justice. Any output of such a machine will only be persuasive because we interpret it as such—but there is at least a risk that a machine that does a reasonable job might be an attractive enough prospect in an overburdened legal system that we gloss over its inherent limitations, opening the door to the kind of problematic reflexive effect we describe below. A further, rather dystopian risk is that the test flips backward on itself, so that human judges start to be tested according to the output of the robot judge—a robotomorphic trajectory that moves toward wholesale datafication of law, in which the assumptions reflected in the design become self-fulfilling prophecies.³¹

The Normativity of Legal Tech Design

We spoke above about the role that automation bias might play in accelerating the adoption of predictive legal tech systems. A further dimension to the attraction of such technologies is the issue of their “technological normativity,” or the ways in which their designs enable and constrain their users’ actions.³²

One need not subscribe to the idea that technological artefacts are political to appreciate that they nevertheless have normativities, or “intentionalities,” that incline them toward particular discrete uses or forms of

²⁸ F. Cabitza, “Breeding Electric Zebras in the Fields of Medicine” [2017] *arXiv preprint arXiv:1701.04077* (discussing the challenges of overriding the predictions of statistical models in the medical domain).

²⁹ Cf. H.S. Sætra, “Robotomorphy” [2021] *AI and Ethics*.

³⁰ E. Volokh, “Chief Justice Robots” [2018] 68 *Duke Law Journal* 1135.

³¹ Sætra, *supra* note 29 at 7. This phenomenon is sometimes referred to as Goodhart’s Law, articulated by Strathern thus: “When a measure becomes a target, it ceases to be a good measure.” See M. Strathern, “‘Improving Ratings’: Audit in the British University System” (1997) 5 *European Review* 305 at 3.

³² On the pivotal distinction between normativities, see M. Hildebrandt, “Legal and Technological Normativity: More (and Less) than Twin Sisters” (2008) 12 *Techné: Research in Philosophy and Technology* 169.

interaction, which in turn affect the broader practices in which they are embedded.³³ This is as true of legal tech systems as it is of any other technology, and so the normativities produced by the design—and in some cases the politics these reflect—will affect the shape of the practice of law, in one way or another. It is, of course, possible to argue, as the Critical Legal Studies (CLS) school has for decades, that law is in any event shaped by politics. But the focus here is different: To what extent do the *underlying* technologies of legal practice have an effect on its normative backdrop? This is a question that comes *before* the analysis of the structures of power reflected in the legal system, when it is viewed purely as a social phenomenon.

Change in legal tech is not new. Just as the late 1990s saw the dawn of the era of the search engine in response to the Web, we have seen a revolution in the tools used in the practice of law over the course of the past several decades. Through the eyes of the lawyer trained in the mid-twentieth century, the introduction of electronic legal databases in the 1980s was no doubt a magical, and for some perhaps a suspicious, development. Few would argue now, however, that such technologies are not improvements on the manual processes that went before, all things being equal.³⁴ So, just as electronic databases became an integral part of the lawyer's practice toward the end of the twentieth century, we might also find that some of the technologies currently being marketed aggressively by legal tech providers have come to be an entirely standard part of the landscape of legal practice.

This is particularly likely with respect to systems that appear to be a relatively incremental advance on what has gone before. An obvious example is legal search, where natural language processing transformer models aim to improve results over standard keyword searching by providing sensitivity to the context within which the search term appears.³⁵ There, the results don't *look* any different; the legal tech application appears to the lawyer more-or-less identical to how it did before—but the underpinnings or back end of the system are entirely different, and the results are derived in a totally different way. The next step in the evolution of such approaches is to frame the statistical prediction in terms of the system "understanding the concepts" that underlie a complex search term, as for example in the case of Casetext's Parallel Search.³⁶ Again, the interface does not change much, but the underlying mechanism for deriving the results that the lawyer sees and acts upon is profoundly different.

³³ L. Winner, "Do Artifacts Have Politics?" [1980] *Daedalus* 121; D. Ihde, *Technology and the Lifeworld: From Garden to Earth* (Indiana University Press, 1990) at 141–142; M. Kranzberg, "'Technology and History': Kranzberg's Laws" (1986) 27 *Technology and Culture* 544

³⁴ Although in some parts of the world, the challenge of democratizing access to legal materials remains.

³⁵ See for example I. Chalkidis et al, "LEGAL-BERT: The Muppets Straight out of Law School," *Findings of the Association for Computational Linguistics: EMNLP 2020* (online: Association for Computational Linguistics, 2020) <<https://aclanthology.org/2020.findings-emnlp.261>> accessed 31 October 2021.

³⁶ "Parallel Search" (Casetext) <<https://casetext.com/parallel-search/>> accessed 31 October 2021.

The Litigator's Practice as a Constitutive Element of Justice

Legal systems bestow certain capacities upon litigators, such as the right to initiate a case before the court, lodge motions and call witnesses. At the same time, they are subject to ethical and professional duties to the courts, their clients and others.³⁷ They have professional obligations to uphold the Rule of Law and maintain the administration of justice.³⁸ What does it mean if, in the course of exercising those capacities and duties, the normative commitment to justice that lies at the core of their profession is, to whatever extent, modulated or supplanted by outputs that are derived solely from statistical induction? The boundary between technological and legal normativity shifts: That part of the practice that is constituted by the legal tech system becomes larger, while the part that exists within the institutional domain shrinks, along with the role of the commitments and safeguards that exist there. While the introduction of tools like word processors and technologies like email had a minimal effect on the normative core of law, we are now faced with a set of technologies whose normativity has the potential to supplant fundamental aspects of the role played by litigators within the broader legal enterprise. Central elements of the litigator's practice—the creative marshalling of facts and legal sources, within the constraints of a normatively guided practice—come to be replaced with outputs that are governed by no more than statistical regularity.

The litigator's practice is one part of the idea that justice must not only be done, but must also be seen to be done.³⁹ Although this idea is usually associated with procedural and evidential propriety, we suggest that it can be extended to encompass the parts of the process that ought to be the product of human capacities, rather than computation or statistical prediction. What must be "seen" are the *parts* of the process as much as the whole, since each can only be understood in light of the other (echoing the notion of the "hermeneutic circle").⁴⁰ If parts of what allows us to characterize the process of argumentation as *legal* are outsourced to machines that have no comprehension of that ideal, there is a risk that the end result has in some sense been compromised, even if it looks acceptable on the face of it. This is as true of the (technological) tools that are used as it is of the procedural steps of the process itself. There is therefore a strong argument to be made that a

³⁷ Paterson, *supra* note 5.

³⁸ See for example, "Ethics in Law" (Law Society of England and Wales) <<https://www.lawsociety.org.uk/topics/regulation/ethics-in-law>> accessed 1 November 2021; "Regulation and Compliance" (Law Society of Scotland) <<https://www.lawscot.org.uk/members/regulation-and-compliance/>> accessed 1 November 2021.

³⁹ To paraphrase the English High Court case of *R. v Sussex Justices ex parte McCarthy* [1924] 1 K.B. 256 at 259, *per* Hewart LCJ: "Justice should not only be done, but should manifestly and undoubtedly be seen to be done." On the centrality of the role of procedure to the Rule of Law, see Waldron, *supra* note 13.

⁴⁰ MacCormick, *supra* note 14 at 48, and on the general concept see H.-G. Gadamer, *Truth and Method* (Joel Weinsheimer and Donald G Marshall trs, London: Bloomsbury, 2013) ch. 4.

commitment to the concept of law we have outlined requires transparency with respect to the kinds of systems that are used in legal practice.

Looked at this way, the productive or creative aspect of legal argumentation is a means of both communicating and legitimating argument within the institutional sphere of litigation—something legal tech can never do on its own. In setting out their legal story, the litigator actively *communicates* one competing vision of justice in the instant case.⁴¹ In a constitutional democracy, this communication serves two symbolic purposes. First, the process of justifying ourselves against some normative yardstick is what makes us responsible—response-able; able to respond—and through this recognition of our agency, we are respected as participants in the social enterprise of the law, even where the outcome is not in our favour.⁴² Second, the public is in principle able to follow that justification at the trial or hearing, to hear the legal premises being relied upon, and how the litigator's argument relates them to the facts as they have been asserted. This is true even for those who disagree with the interpretation of the facts or norms that have been put forward; the point is that the visibility of the argument and of its component parts legitimizes the outcome, even for those who disagree.

Thus, not just any process will do—the elements of a valid process are those that make real the values of the Rule of Law. Lawyers do not simply “go through the motions,” but rather apply their creative capacities in the service of the overall enterprise. Without appropriate opportunity to exercise that creativity in service of a normative end, the notion of justice reflected in the result of the process will risk being in some sense flattened or stunted. To the extent that Holmes's prediction is coming true—that the lawyer “of the future is the man of statistics”⁴³—we lose something profound at the heart of what enables the law to be just. If elements of the litigator's practice are “black boxed” by legal tech systems whose design reflects something other than the will to justice, the communicative and legitimizing role of the legal process is necessarily manipulated, to whatever degree.

Institutionality and Determinism

At the extreme end of the spectrum, this scenario is problematic from the perspective of law as an institutional order characteristic of constitutional states.⁴⁴ Such states rely on the existence of a predefined process of public argumentation that can authoritatively determine, within this shared institutional understanding, or legal “Welt,”⁴⁵ the outcome of a dispute about

⁴¹ Cf. E. Esposito, “Transparency vs. Explanation: The Role of Ambiguity in Legal AI” (2021) 1 *Journal of Cross-disciplinary Research in Computational Law*.

⁴² J. Gardner, “The Mark of Responsibility” (2003) 23 *Oxford Journal of Legal Studies* 15. See also the discussion of Waldron's account of the dignitarian concept of law above.

⁴³ Holmes, *supra* note 19.

⁴⁴ N. MacCormick, *Institutions of Law: An Essay in Legal Theory* (New York: Oxford University Press, 2007) ch. 1–3.

⁴⁵ Hildebrandt, *Smart Technologies*, *supra* note 9 at 53–54.

rights, duties, and liabilities.⁴⁶ These do not exist independently in our minds as objects of thought. Instead, they are continually recognized through the combination of externalized texts and the human faculty of attributing (legal) meaning, each considered in the light of whatever future context this recognition takes place.⁴⁷

If the lawyer was simply a manipulator of legal signs, akin to a computer shifting bits according to a predefined programme, then indeed the goal of legal practice would in effect be to identify the “programme”—to infer it from the heap of potentially applicable legal norms. Alternatively, if the legal programme is induced by the machine from the data available to it, and the lawyer simply takes this output as the central postulate of their argument, they are potentially again taking an output, or a computational “reading,” as the defining pillar of the legal story they are tasked with constructing.

The question is what it means if the lawyer comes to be treated as merely a user of such systems, taking the output as-is and importing it into the legal process, and relying on it as a statement of, or as best reflecting, the law. Can their “translation” of a statistical prediction into the institutional world of law, possible because of their status as lawyer, render it institutionally valid, even where they do not fully understand the statistical mechanisms behind its production?

The Necessity of “Lossless” Law

The outputs of legal tech systems may gain a kind of legitimacy by dint of the legal community treating them as valid. The predictions would thus make their way into the corpus of legal text, either directly (as in the case of machine-generated motions, arguments or pleadings) or indirectly by shaping the sources or arguments that are relied upon in individual court proceedings. Again, this is something that is likely to change slowly over time, as more lawyers rely on these systems and so the predictions start to be reflected in the corpus of legal text through their use in the courtroom.⁴⁸ This could result in a feedback loop, where statistically promoted results can gain further prominence independently of their legal relevance, the effect increasing over the course of several cycles. Such prominence would be determined solely on the basis of statistical proxies for relevance, similarity, or performance more generally.⁴⁹ This will be true of legal search, where particular cases might be given undue

⁴⁶ Kerr warns of the “potentially deep systemic problems sure to arise as we scuttle the justice system in favour of efficient actuarial models.” Kerr is speaking here about prediction of the behaviour of citizens rather than the behaviour of the courts, but he warns more generally of the implications of predictive technologies. A predictive approach, Kerr argues, is linked to a society organized around risk. See Kerr, *supra* note 5 at 26.

⁴⁷ Hildebrandt, *Smart Technologies*, *supra* note 9 at 53–54.

⁴⁸ For an empirical study that demonstrates the differences in results between various legal search providers (Casetext, Fastcase, Google Scholar, Lexis Advance, Ravel, and Westlaw, as well as Google Scholar), see S.N. Mart, “The Algorithm as a Human Artifact: Implications for Legal {Re}Search” (2017) 109 *Law Library Journal* 387.

⁴⁹ On the fundamental distinction of human and statistical “truth,” see D. McQuillan, “Data Science as Machinic Neoplatonism” (2018) 31 *Philosophy & Technology* 253.

prominence owing to their fit with a statistical pattern,⁵⁰ and with systems that generate suggested texts, where those end up reflected in the documents associated with the case (as in the briefs published in the United States).⁵¹

One can imagine in this process a notional point of inflection, at which lawyers start to adapt their real-world practices to reflect those predictions⁵² (recall the discussion above of robotomorphy), with the openness of interpretative possibilities beginning to narrow.⁵³ At that moment, practice and prediction start to converge, the latter distorting the former according to what is *statistically optimal*,⁵⁴ rather than in accordance with any other guiding normative value.⁵⁵

The output of a statistically driven legal search or motion writing system will necessarily impose some technological abstraction, selection, or representation on the linguistic “data” before the point at which the law, through lawyers, can construct institutional facts that are unencumbered by that framing. The results of the search or the proposed text are ranked according to the performance metric designed into the system, rather than any normatively directed notion of legal relevance. The law is thus somewhat displaced, if the underlying medium has already framed the materials of practice before the lawyer has had a chance to take the initial step of rendering text or other evidence into a form tractable to legal argumentation.⁵⁶ The outputs derived from the data, and even the framing of legal text as “data” in the first place, mean that instead of the “raw texts” of law being the basis of the lawyer’s practice, a technologically and statistically mediated representation is used instead. This is the potential effect of the mediation of legal text, *qua* data, by legal technologies; instead of dealing directly with “lossless” law, that is the raw texts that are its primary manifestation, we instead work with “lossy” law, a “compressed” version

⁵⁰ For a discussion of an analogous effect in the results provided by Google Scholar, see C. Rovira, L. Codina and C. Lopezosa, “Language Bias in the Google Scholar Ranking Algorithm” (2021) 13 *Future Internet* 31.

⁵¹ The Supreme Court of the United States offers a guide on how to access Supreme Court briefs. See “Where to Find Briefs of the Supreme Court of the U.S.” (*Supreme Court of the United States*) <<https://www.supremecourt.gov/meritsbriefs/briefsource.aspx>> accessed 1 November 2021. Casetext’s “Compose” feature allows the automated drafting of briefs. See E. Cheek, “Which Types of Briefs Can You Draft Using Compose?” <<https://help.casetext.com/en/articles/4044930-which-types-of-briefs-can-you-draft-using-compose>> accessed 1 November 2021.

⁵² As Rovira et al note, academics may be under pressure to perform “academic search engine optimization,” to ensure their work is visible in Google Scholar search rankings. See Rovira, Codina and Lopezosa, *supra* note 50.

⁵³ L. Diver, “Normative Shortcuts and the Hermeneutic Singularity” <<https://www.cohubicol.com/blog/normative-shortcuts-and-the-hermeneutic-singularity>> accessed 1 November 2021. See also J. Cobbe, “Legal Singularity and the Reflexivity of Law” [2020] *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence* 107.

⁵⁴ Mitchell’s canonical definition of machine learning is useful here: “A computer program is said to learn from experience *E* with respect to some class of tasks *T* and performance measure *P*, if its performance at tasks in *T*, as measured by *P*, improves with experience *E*.” See T.M. Mitchell, *Machine Learning* (New York: McGraw-Hill, 1997) at 2.

⁵⁵ For a useful discussion of the manipulative force of optimization in the context of health apps, see M. Sax, “Optimization of What? For-Profit Health Apps as Manipulative Digital Environments” [2021] *Ethics and Information Technology* (“One would expect that health apps are aimed at optimizing *the health* of their users, but in reality the need to monetize their userbase leads for-profit health apps to, first and foremost, optimize user engagement and, in effect, conversion of users” [p. 2, emphasis supplied, references omitted].)

⁵⁶ Cf. van Dijk, *supra* note 18 at 162, discussing how lawyers translate “a whole range of heterogeneous phenomena” into the world of law, enabling them to “speak legally.”

derived from the original that might look and feel the same but, through the interpolation of a data-driven frame, has necessarily refracted away some of its original fidelity—and in ways the lawyer might not be aware of.⁵⁷ In other contexts this might not matter—we tend not to be too concerned about less-than-perfect search results on the Web, for example⁵⁸—but in law these statistically mediated representations can in effect become the substance of law, if they are unwittingly trusted as such.⁵⁹

The Determinisms of Algorithmic and Interface Design

Viewed this way, and echoing the distinction drawn above, there are two ways to look at this overdetermination of the lawyer's practice.⁶⁰ First is a kind of technological determinism: The lawyer becomes in effect an operator, or "user," of the machine, forming a bridge between its predictions (both in the technical sense of machine learning outputs and the sense of predicting legal outcomes) and the concrete legal process, by inserting the former into the procedural steps of the latter. The output of the legal tech becomes part of the law, whether it takes the form of search results that prioritize case A over case B,⁶¹ or a case law analytics platform that predicts the success of a particular verbal construction in a motion, both of which operate only on the basis of statistical regularities in past data. At an extreme, we can modify Holmes's original prediction: law becomes "what the *legal tech* will do in fact, and nothing more pretentious."⁶²

This might seem far-fetched, and we hope so. But the second notion of overdetermination is subtler and for that reason potentially more problematic from a democratic perspective. Here, the computational outputs do not *determine* what the lawyer argues in court, but instead subtly shift it this way or that, introducing reflexive effects that might be difficult at first to detect. The effects in question are not instances of "technological management,"⁶³ where the obvious circumscription of action attracts greater scrutiny because of the clear possibility that their normative power might be abused. Societies are rightly concerned about automated decision-making systems that enforce certain undesirable or unacceptable biases, particularly around race, gender,

⁵⁷ The analogy here is, of course, with lossless and lossy image and audio formats, such as PNG and JPEG, and FLAC and MP3, respectively, where in the latter data is discarded and/or compressed to produce a version that is "good enough" for its target audience, in terms of the trade-off between quality and file size.

⁵⁸ But see the discussion of biased Google Scholar results in Rovira, Codina and Lopezosa, *supra* note 53.

⁵⁹ Cabitza refers to the biasing effect of computational framing in the medical domain as "empirical anopsia," which "occurs when there is a lack of awareness of the essential arbitrariness of the aspects that any data structure and its values depict of a real and continuous phenomenon, as well as of those elements of this phenomenon that any of its representations amplifies, conceals, or worse yet (necessarily), distorts." See Cabitza, *supra* note 28 at 6.

⁶⁰ Cf. M. Hildebrandt, "Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning" (2019) *20 Theoretical Inquiries in Law*, discussing the role of data protection legislation in protecting against choice architectures that diminish data subjects' agency.

⁶¹ Mart's research demonstrates that different search algorithms produce very different results. Mart, *supra* note 48 (comparing the ten top-ranked search results generated by Casetext, Fastcase, Google Scholar, Lexis Advance, Ravel, and Westlaw in relation to the same search query).

⁶² Cf. Holmes, *supra* note 19 at 461.

⁶³ R. Brownsword, "Technological Management and the Rule of Law" (2016) *8 Law, Innovation and Technology* 100. See also L. Diver, *Digisprudence: Code as Law Rebooted* (Edinburgh: Edinburgh University Press, 2022).

sexuality, and socio-economic status, and the scrutiny directed at such systems is welcome in exposing the inequalities to which they contribute. But the kinds of bias we are concerned with here are not quite so apparent, given that they do not necessarily reflect an existing societal prejudice that can be identified in the outcome of the decision. The lack of immediate signals is precisely why the effects of such technologies are potentially more problematic in terms of normative legal theory, since any instinctive aversion to illegitimate technological management is not triggered; the reflexive effects of certain categories of legal technology are subtle enough that they are likely to accumulate slowly over time, shifting the bedrock of legal practice in ways that may not be apparent at any given moment.

In statistical legal tech systems, the effects in question operate at a lower level, affecting the very conditions of possibility for argument—and decision-making. The technological normativity of the statistically governed output, and of the way in which its interface presents it to the lawyer, impresses on them one particular view of the law that is framed by the design of that system. To the extent that the output is assumed to be “better,” the lawyer’s reasoned interpretation of legal texts is moderated, and with it their responsiveness to the animating role of ambiguity in bringing legal contestability to life.⁶⁴ At the most extreme end, the pattern-based prediction *becomes* the argument, without further ado—the lawyer simply presents the argument that the machine says will most likely be effective. The interface of the system—its particular layout and ways of presenting the statistical result—becomes, through the lawyer’s translation of it into the legal process, in effect a statement of the content of the law, especially if it is then accepted by the court.

A legal tech system that is marketed as particularly effective or accurate is all the more likely to have such an effect. We are confronted not with technological *determinism*, but with the guiding effects of technological *normativity*, where the action-enabling and constraining aspects of the system’s design have a guiding effect on the path of the law, through successively nudged usages that themselves feed back into the underlying corpus of legal text.

Natural language has always and necessarily required interpretative work,⁶⁵ highlighting the requirement that the lawyer posit a reasoned argument in favour of a given interpretation that protects their client’s interests and upholds justice. The normative effect of a legal tech interface that provides quick answers, without the need for the interrogation of text(s) necessary to foment true legal reasoning, will potentially flatten this constitutive practice. Statistical prediction makes easy what natural language had previously ensured was (productively) difficult.

⁶⁴ Esposito, *supra* note 41.

⁶⁵ Cf. Gadamer, *supra* note 40.

Is Vigilance Enough?

We noted above that lawyers have professional duties to their clients and the courts. Respect for the court requires that the lawyers should be familiar with the facts of the case and the relevant law.⁶⁶ In the event that a lawyer relied exclusively on the output of legal tech, they might find themselves in breach of their professional obligations and their duties to the court. Commercial considerations may also come into play. Legal tech providers are generally careful to include suitable disclaimers in their terms and conditions. Jus Mundi's Terms of Use, for example, show careful attention both to the limits of their offerings and the responsibilities of lawyers.⁶⁷ These constraints may militate against wholesale reliance on the output of legal tech. However, as noted earlier, in all likelihood the real risk lies not so much in wholesale reliance, but in the subtle reshaping of the lawyer's arguments and strategy. Is it possible to preclude or counter this effect? Can we rely on lawyers' vigilance to prevent it?

It is not obvious that we can. True, a lawyer might stolidly resist the more overblown claims of legal tech marketing.⁶⁸ At the same time, the best legal tech providers are open and honest about the choices embedded in the architecture of their systems;⁶⁹ in those cases a lawyer can inform themselves about the fact that such choices have been made, even if they do not fully understand the implications. Legal education might be broadened to educate future lawyers about the design choices and the capabilities and limitations of legal tech. It remains unclear, however, whether these measures are sufficient to ensure that the output of legal tech, with its inherent capabilities and limitations, does not shape both the content and the fabric of law. In theory, one could mitigate this risk by using legal tech only in parallel with traditional methods. For example, lawyers might supplement their searches on platforms like Westlaw or Lexis Nexis with research using other resources: commercial search engines, specialist publications or blogs, and, of course, their own domain knowledge. But this is precisely what legal tech providers claim is no longer necessary.

⁶⁶ See, for example, R. Bell and C. Abela, "A Lawyer's Duty to the Court," *Proceedings of a Symposium on Professionalism* (2012) <http://advokat-id.ru/wp-content/uploads/2018/06/11024_10167_CEA-A-Lawyers-Duty-to-the-Court.pdf> accessed 1 November 2021.

⁶⁷ "You are solely responsible for the consultation, choice, use and interpretation of the documentation provided by Jus Mundi through the Website, as well as for your actions and advice provided in the context of your professional practice. You recognize that the Services offer you an additional solution but not an alternative to the means you already use and that the Services do not replace them." See "General Terms of Use" (*Jus Mundi*) <<https://jusmundi.com/en/terms-of-use>> accessed 1 November 2021.

⁶⁸ O'Grady notes with surprise that the readers of her blog did not seem to be especially "wowed" by the AI powered search functionality of Westlaw Edge. See J. O'Grady, "Hits and Misses Part 4: Westlaw Edge—Hit, Miss or Hold Off? Customers Respond—Show Me the ROI!" <<https://www.deweybstrategic.com/2019/05/hits-and-misses-part-4-westlaw-edge-hit-miss-or-hold-off-customers-respond-show-me-the-roi.html>> accessed 1 November 2021.

⁶⁹ Casetext, for example, provides an accessible account of how they try to improve their relevance rankings. See "Search Results Evaluation Efforts at Casetext" <<https://casetext.com/blog/search-results-evaluation-efforts-at-casetext/>> accessed 1 November 2021.

In practice, it is difficult to know to what extent the output of legal tech nevertheless shapes the lawyer's assessment of the law and the strategy to be adopted in a particular case. At the point of interaction, it might be argued that the inherent presence of a "human in the loop" is sufficient to mitigate the effect of any especially off-kilter outputs. This kind of mitigation is usually aimed at automated decision-making systems that produce legal effects.⁷⁰ But having a "lawyer in the loop" is not of much assistance for a great many legal tech applications, such as the legal search and strategy support systems we have been discussing, where the effect of the prediction at the point it is made does not entail any direct change in legal status. Furthermore, although in principle the lawyer has the chance to disregard the output, to place the onus on them to overrule the system is to ignore the problems of technological normativity described above. The notional overseeing "human in the loop" in such a case would be precisely the person who is subjected to the normative effects we have been discussing.

Conclusion

Taking a data-driven shortcut is attractive for many reasons: reduced cost, faster results, and greater throughput in the justice system. But these are tempting sirens that can easily distract us from the need to protect the creative linguistic spaces that are constitutive of legal argumentation, and that facilitate the kinds of communication that are a precondition for the acceptability of law.

Our intuition is that something of the process of argumentation is altered or lost by interpolating machine output between the client's explanation of their position, and the lawyer's expression of that position in court. In the end, legal tech that generates drafts of documents proceeds either on the basis of encoded rules, or on the basis of statistical analysis. It neither listens to the client nor has any (semantic/normative) understanding of the law.⁷¹ Similarly, though less directly, advanced legal search systems also affect the litigator's presentation of the argument, by providing the materials for that argument according to a purely statistical notion of relevance. In both instances there is a problematic fit with the broader normative and historical context of the legal order.

It might be argued that this does not matter, since the legal tech is merely an aid, a tool that assists but does not supplant the lawyer. We do not think this argument stands up to scrutiny. It is often said that "We shape our tools and thereafter our tools shape us."⁷² Use of these systems, we suggest,

⁷⁰ Cf. Article 22 of the General Data Protection Regulation (GDPR).

⁷¹ Its output cannot be compared to that of an inexperienced or junior lawyer. The junior lawyer may make mistakes, but they possess both an ability to listen and a sense of justice or fairness which the machine lacks.

⁷² This aphorism is very often attributed to Marshall McLuhan, but see A. Kuskis, "We Shape Our Tools and Thereafter Our Tools Shape Us" <<https://mcluhangalaxy.wordpress.com/2013/04/01/we-shape-our-tools-and-thereafter-our-tools-shape-us/>>

conditions us to think that they deliver “better.” Yet, without the ability to listen to a client or absorb the meaning of legal norms and principles, the “intelligence” that such a system brings to its task is limited. Its assessment of whether one case is similar to another can only be based on finite rules or on more or less sophisticated pattern matching. It cannot assess whether cases are similar by analogy, whether a generic (not domain-specific) principle is at stake, or whether fairness in the particular case might support an argument and an outcome that, from the perspective of existing case law, might look like an outlier or non-starter.

The combination of the promise of these systems (which are, after all, marketed on the basis that they deliver “better”) and the limitations of their output may effectively mean that a “third voice”—neither that of the client, nor that of the law—is introduced into the argumentation process. We think this is problematic, and may have implications—not just for the instant case but, more importantly, for the normative structure of law. The onus is on the legal profession and the providers of legal tech to show the contrary. Law cannot be captured in its entirety by rules and statistics. Lawyers should not be persuaded that it can. People must be capable of arguing for and securing change in the law. The “dignitarian” aspects of legal argumentation require that lawyers should be open to that possibility.

tools-shape-us/> accessed 31 October 2021 (suggesting that the quote should be attributed to Fr. John Culkin SJ, a friend of McLuhan).

References

Bańkowski Z, "Don't Think About It: Legalism and Legality" in M.M. Karlsson, Ó. Páll Jónsson and E.M. Brynjarsdóttir (eds), *Rechtstheorie: Zeitschrift für Logik, Methodenlehre, Kybernetik und Soziologie des Rechts* (Berlin: Duncker & Humblot, 1993).

Bell R and C Abela, "A Lawyer's Duty to the Court," *Proceedings of a Symposium on Professionalism* (2012) <http://advokat-id.ru/wp-content/uploads/2018/06/11024_10167_CEA-A-Lawyers-Duty-to-the-Court.pdf> accessed 1 November 2021.

Brownsword R, "Technological Management and the Rule of Law" (2016) 8 *Law, Innovation and Technology* 100.

Cabitza F, "Breeding Electric Zebras in the Fields of Medicine" [2017] *arXiv preprint arXiv:1701.04077*.

Cantwell Smith B, *The Promise of Artificial Intelligence: Reckoning and Judgment* (Cambridge, MA: The MIT Press, 2019).

Chalkidis I et al., "LEGAL-BERT: The Muppets Straight out of Law School," *Findings of the Association for Computational Linguistics: EMNLP 2020* (online: Association for Computational Linguistics, 2020) <<https://aclanthology.org/2020.findings-emnlp.261>> accessed 31 October 2021.

Cheek E, "Which Types of Briefs Can You Draft Using Compose?" <<https://help.casetext.com/en/articles/4044930-which-types-of-briefs-can-you-draft-using-compose>> accessed 1 November 2021.

Cobbe J, "Legal Singularity and the Reflexivity of Law" [2020] *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence* 107.

Diver L, "Normative Shortcuts and the Hermeneutic Singularity" <<https://www.cohubicol.com/blog/normative-shortcuts-and-the-hermeneutic-singularity>> accessed 1 November 2021.

—, *Digisprudence: Code as Law Rebooted* (Edinburgh: Edinburgh University Press, 2022).

Esposito E, "Transparency vs. Explanation: The Role of Ambiguity in Legal AI" (2021) 1 *Journal of Cross-disciplinary Research in Computational Law*.

"Ethics in Law" (Law Society of England and Wales) <<https://www.lawsociety.org.uk/topics/regulation/ethics-in-law>> accessed 1 November 2021.

Fish SE, *Doing What Comes Naturally: Change, Rhetoric, and the Practice of Theory in Literary and Legal Studies* (Durham: Duke University Press, 1989).

Gaakeer J, *Judging From Experience: Law, Praxis, Humanities* (Edinburgh University Press, 2019).

Gadamer H-G, *Truth and Method* (Joel Weinsheimer and Donald G Marshall trs, London: Bloomsbury, 2013).

Gardner J, "The Mark of Responsibility" (2003) 23 *Oxford Journal of Legal Studies* 15.

"General Terms of Use" (*Jus Mundi*) <<https://jusmundi.com/en/terms-of-use>> accessed 1 November 2021.

G'sell F, "Predicting Courts' Decisions Is Lawful in France and Will Remain So" <<https://gsell.tech/en/predicting-courts-decisions-is-lawful-in-france-and-will-remain-so/>> accessed 30 September 2021.

Harvey DJ, *Collisions in the Digital Paradigm: Law and Rule Making in the Digital Age* (Bloomsbury, 2017).

Hildebrandt M, "Legal and Technological Normativity: More (and Less) than Twin Sisters" (2008) 12 *Techné: Research in Philosophy and Technology* 169.

—, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (London: Edward Elgar Publishing, 2015).

—, "Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics" (2018) 68 *University of Toronto Law Journal* 12.

—, "Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning" (2019) 20 *Theoretical Inquiries in Law*.

—, "Code-Driven Law: Freezing the Future and Scaling the Past" in S.F. Deakin and C. Markou (eds), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence* (Oxford, UK; New York, NY: Hart, 2020).

Holmes OW, "The Path of the Law" (1897) 10 *Harvard Law Review* 457.

Ihde D, *Technology and the Lifeworld: From Garden to Earth* (Indiana University Press, 1990).

Katz DM, "Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry" (2012) 62 *Emory Law Journal* 909.

Kerr I, "Prediction, Pre-Emption, Presumption: The Path of Law After the Computational Turn" in M. Hildebrandt and K. De Vries (eds), *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of technology* (London: Routledge, 2013).

Kranzberg M, "Technology and History: 'Kranzberg's Laws'" (1986) 27 *Technology and Culture* 544.

Kuskis A, "We Shape Our Tools and Thereafter Our Tools Shape Us" <<https://mcluhangalaxy.wordpress.com/2013/04/01/we-shape-our-tools-and-thereafter-our-tools-shape-us/>> accessed 31 October 2021.

Latour B, "Biography of an Inquiry: On a Book About Modes of Existence" (2013) 43 *Social Studies of Science* 287.

"Lawtech and Ethics Principles" (Law Society of England and Wales, 2021) <<https://www.lawsociety.org.uk/topics/research/lawtech-and-ethics-principles-report-2021>> accessed 30 October 2021.

"Lex Machina: Legal Analytics® for the Data-Driven Lawyer" (Lex Machina) <<https://lexmachina.com/wp-content/uploads/Legal-Analytics-for-Data-Driven-Lawyer.pdf>> accessed 31 October 2021.

Livermore MA and DN Rockmore (eds), *Law as Data: Computation, Text, and the Future of Legal Analysis* (Santa Fe: SFI Press, 2019).

MacCormick N, *Rhetoric and the Rule of Law: A Theory of Legal Reasoning* (Oxford; New York: Oxford University Press, 2005).

—, *Institutions of Law: An Essay in Legal Theory* (New York: Oxford University Press, 2007).

Mart SN, "The Algorithm as a Human Artifact: Implications for Legal {Re}Search" (2017) 109 *Law Library Journal* 387.

Mason S, "Artificial Intelligence: Oh Really? And Why Judges and Lawyers Are Central to the Way We Live Now—but They Don't Know It" [2017] *Artificial intelligence* 29.

McQuillan D, "Data Science as Machinic Neoplatonism" (2018) 31 *Philosophy & Technology* 253.

Meneceur Y and C Barbaro, "Artificial Intelligence and the Judicial Memory: The Great Misunderstanding" [2021] *AI and Ethics*.

Mitchell TM, *Machine Learning* (New York: McGraw-Hill, 1997).

O'Grady J, "Hits and Misses Part 4: Westlaw Edge—Hit, Miss or Hold Off? Customers Respond—Show Me the ROI!" <<https://www.deweybstrategic.com/2019/05/hits-and-misses-part-4-westlaw-edge-hit-miss-or-hold-off-customers-respond-show-me-the-roi.html>> accessed 1 November 2021.

"Parallel Search" (Casetext) <<https://casetext.com/parallel-search/>> accessed 31 October 2021.

Paterson A, "Duties to the Court," *Law, Practice & Conduct for Solicitors* (2nd ed., W Green/Thomson Reuters, 2014).

"Regulation and Compliance" (Law Society of Scotland) <<https://www.lawscot.org.uk/members/regulation-and-compliance/>> accessed 1 November 2021.

Rouvroy A and B Stiegler, "The Digital Regime of Truth: From the Algorithmic Governmentality to a New Rule of Law" (2016) 3 *La Deleuziana: Online Journal of Philosophy* 6.

Rovira C, L Codina and C Lopezosa, "Language Bias in the Google Scholar Ranking Algorithm" (2021) 13 *Future Internet* 31.

Sætra HS, "Robotomorphy" [2021] *AI and Ethics*.

Sax M, "Optimization of What? For-Profit Health Apps as Manipulative Digital Environments" [2021] *Ethics and Information Technology*.

"Search Results Evaluation Efforts at Casetext" <<https://casetext.com/blog/search-results-evaluation-efforts-at-casetext/>> accessed 1 November 2021.

Strathern M, "'Improving Ratings': Audit in the British University System" (1997) 5 *European review* 305.

"THE FUTURE OF LAW Judge Analytics: Predicting the Behavior of the Courtroom's Boss" (*Lex Machina*) <https://pages.lexmachina.com/Website_Future-of-Law-Webcast_The-Future-of-Law-3.html> accessed 31 October 2021.

van Dijk N, "The Life and Deaths of a Dispute: An Inquiry Into Matters of Law" in K. McGee (ed.), *Latour and the Passage of Law* (Edinburgh University Press, 2015).

Volokh E, "Chief Justice Robots" (2018) 68 *Duke Law Journal* 1135.

Waddington M, "Research Note. Rules as Code" (2020) 37 *Law in Context. A Socio-legal Journal* 1.

Waldron J, "The Rule of Law and the Importance of Procedure" (2011) 50 *Nomos* 3.

"Where to Find Briefs of the Supreme Court of the U.S." (*Supreme Court of the United States*) <<https://www.supremecourt.gov/meritsbriefs/briefsource.aspx>> accessed 1 November 2021.

Winner L, "Do Artifacts Have Politics?" [1980] *Daedalus* 121.