

High Tech, Low Fidelity? Statistical Legal Tech and the Rule of Law

Diver, Laurence; McBride, Pauline

Published in:
Verfassungsblog

DOI:
[10.17176/20220401-131228-0](https://doi.org/10.17176/20220401-131228-0)

Publication date:
2022

License:
CC BY-SA

Document Version:
Final published version

[Link to publication](#)

Citation for published version (APA):
Diver, L., & McBride, P. (2022). High Tech, Low Fidelity? Statistical Legal Tech and the Rule of Law. *Verfassungsblog*. <https://doi.org/10.17176/20220401-131228-0>

Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

Take down policy

If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

High Tech, Low Fidelity? Statistical Legal Tech and the Rule of Law

VB verfassungsblog.de/roa-high-tech-low-fidelity/



Laurence Diver



Pauline McBride

This article belongs to the debate » [The Rule of Law versus the Rule of the Algorithm](#)
01 April 2022

The advent of statistical ‘legal tech’ raises questions about the future of law and legal practice. While it has always been the case that technologies have 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, particularly that of litigators. These applications mediate how practicing lawyers interact with the legal system. By shaping how law is ‘done’, the applications ultimately come to shape what law *is*.

The European Commission’s proposal for a new AI Act hints at what is at stake. It 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.”

We think the concern ought to extend beyond judicial authorities, however. Statistical legal tech may impact on certain fundamentally creative aspects of the litigator’s role, aspects that are central to her professional and ethical duty to respond appropriately to the unique circumstances of her client’s case. Proper exercise of that duty is essential to the Rule of Law; legal practitioners have a normative commitment to legal protection that must be manifested through their creative marshalling and synthesis of legal and factual resources. Any technology whose normative choice architectures might collide with that duty must be closely scrutinised before (and after) it is put into use.

The Purpose of Creative Argumentation

The role of lawyers is to act as agents for those they represent. Their primary duty is to their client, though they also owe duties to the courts and to their fellow lawyers. The motions, pleadings, arguments are their *client’s* motions, pleadings, arguments. The function of law is to enable such argumentation, allowing 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.”² What matters is the process of argumentation. As

Waldron argues, it is this which ensures “respect for the freedom and dignity of each person as an active intelligence.”³⁾ The decisions of courts, therefore, should be regarded not as mere win/lose ‘outputs’, but as the results of a process that involves listening to people’s views, while also constraining them to formulate those views against the backdrop of relevant legal norms.

There is no platonic, logically deducible ‘legal truth’ that already exists ‘out there’, ready to be found by either the lawyer or by a machine. Instead, the creative work of interpretation and argumentation undertaken by the litigator is exactly what is entailed in *creating* legal truth, through the process of ‘doing law’. It is only in presenting an argument, creatively synthesised from its constituent parts, that law thus is *performed*, and Rule of Law values sustained.

Statistically Mediated Sources of Law

The introduction of statistical legal tech poses a subtle but fundamental risk to creative argumentation. If the outputs of legal tech systems built round machine learning prediction are treated as a definitive statement of the law, lawyers may not be *replaced*, but their practices may be *altered*, in ways that affect the normative texture of law.

The Lure of Effectiveness

Change in legal tech is of course not new. We have seen a revolution in the tools used in the practice of law over the past several decades, and just as electronic databases became an integral part of the lawyer’s practice toward the end of the twentieth century, so too might we find that some of the legal technologies being aggressively marketed at present will come to be a standard part of the legal landscape. This is particularly likely for systems that appear to be a relatively incremental advance on what has gone before. In legal search, for example, standard keyword search is being replaced or augmented with contextual search which uses natural language processing techniques to optimise search results through sensitivity to the context within which a search term appears. The results are presented more-or-less exactly as they were before – but the underpinnings or back-end of the system are entirely different, and those results derived in a totally different way.

Commercial legal tech providers often make impressive assertions about the performance and capabilities of their systems. While it might reasonably be argued that most lawyers’ intuitions will insulate them from some of the more sensational claims that are made, the combination of network effects and automation bias is compelling, especially when such systems purport to reduce costs and increase efficiency. Moreover, in most cases these systems *actually work* – in the sense that they generate outputs that are intelligible and usable. Paradoxically, this is precisely where the risk is greatest. The apparent success of the statistical output, bolstered by its speed and by claims made by its provider about the numbers and kinds of variables that factor into its inductive reasoning, creates the sense that the system 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 she can, the conscientious lawyer might even come to believe she *must* use such systems – especially if professional guidance is silent on the broader and deeper implications of their adoption.

A lawyer who relied exclusively on the output of legal tech might find herself in breach of her professional obligations and her duties to the court. The real risk, however, 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 obviate it?

It is not obvious that we can. It is true that a lawyer might resist the more overblown claims of legal tech marketing. The best legal tech providers are open and honest about the choices embedded in the architecture of their systems; in those cases a lawyer can inform herself about the fact that such choices have been made, even if she does 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. In principle one could mitigate this risk by using legal tech only in parallel with traditional methods – but this would be extremely inefficient, and is precisely what legal tech providers claim is no longer necessary. It thus remains unclear what would be required to protect against potentially damaging reshaping of both the content and the fabric of law.

In the long-term, we might start to see in statistical legal tech an idealised 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 proficiency, we try to 'optimise' ourselves and our practices to align with that ideal.

The Importance of 'Lossless' Law

If adopted unthinkingly, the outputs of statistical legal tech systems might gain a kind of legitimacy simply by virtue of the legal community treating them as valid. Over time, the predictions make their way into a jurisdiction's 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. This may result in a feedback loop, where statistically-promoted results gain further prominence independently of their legal relevance. Such prominence would be determined solely on the basis of statistical proxies for relevance, similarity, or performance more generally.

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

This is the potential effect of the mediation of legal text by statistical 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 filtered from the original that appears plausible when presented through the interface of a legal tech application but which, through the interpolation of a data-driven frame, has necessarily lost some of the source’s original fidelity – and in ways unlikely to be apparent to the lawyer.

To Maintain Fidelity, We Must Preserve Creativity

Taking a data-driven approach is attractive for many reasons: reduced cost, faster results, and greater throughput in the justice system. But these are temptations that can easily distract us from the need to protect the creativity that is key to legal argumentation, and that facilitates the kinds of contestation that is a precondition for the Rule of Law.

Something of the nature and quality of argumentation is 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, statistical legal tech that, for example, generates drafts of documents or outputs search results neither listens to the client nor has any (semantic/normative) understanding of the law. It offers up raw materials for the lawyer’s arguments according to a purely statistical notion of what is relevant. In both instances there is a potentially 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 in the hands of 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”. Use of these systems, we suggest, 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.

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. 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 requires that lawyers should be open to that possibility.

The authors are part of the ,Counting as a Human Being in the Era of Computational Law‘ (COHUBICOL) research project.

References

↑1 Proposal for an Artificial Intelligence Act (COM(2021) 206 final), Annex III, Article 8.

↑2 J. Waldron, 'The Rule of Law and the Importance of Procedure' (2011) 50 *Nomos* 3, p. 19.

↑3 *Ibid.*, p. 23.

↑4 See for example Jake Heller, 'What a difference a few years makes: The rapid change of legal search technology' (available at <<https://blog.casetext.com/what-a-difference-a-few-years-makes-the-rapid-change-of-legal-search-technology-ce91a60ab73b>> (accessed 7 March 2022).

↑5 For example, the Law Society of England and Wales' refer to 'rule of law' 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 Principles' (Law Society of England and Wales, 2021) pp. 13–14 <<https://www.lawsociety.org.uk/topics/research/lawtech-and-ethics-principles-report-2021>> accessed 30 October 2021.

↑6 H.S. Sætra, 'Robotomorphy' [2021] *AI and Ethics*.

↑7 Mitchell's 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) p. 2.

↑8 The analogy here is of course with lossless and lossy image and audio formats, such as PNG and JPEG, and FLAC and MP3.

↑9 Its output cannot be compared to the work of an inexperienced or junior lawyer. The latter may make mistakes, but she possesses both an ability to listen and a sense of justice or fairness which the machine lacks.

References

LICENSED UNDER CC BY SA

SUGGESTED CITATION Diver, Laurence; McBride, Pauline: *High Tech, Low Fidelity? Statistical Legal Tech and the Rule of Law*, *VerfBlog*, 2022/4/01, <https://verfassungsblog.de/roa-high-tech-low-fidelity/>, DOI: [10.17176/20220401-131228-0](https://doi.org/10.17176/20220401-131228-0).

Explore posts related to this:

LICENSED UNDER CC BY SA