Law as Computation in the Era of Artificial Legal Intelligence

Hildebrandt, Mireille

Published in:
University of Toronto Law Journal

DOI:
10.3138/utlj.2017-0044

Publication date:
2018

Document Version:
Accepted author manuscript

Draft pre-edited paper for special issue of University of Toronto Law Journal

Law *as* computation in the era of artificial legal intelligence Speaking law to the power of statistics

Mireille Hildebrandt

Abstract

The idea of artificial legal intelligence stems from a previous wave of artificial intelligence, then called jurimetrics. It was based on an algorithmic understanding of law, celebrating logic as the sole ingredient for proper legal argumentation. However, as Holmes noted, the life of the law is experience rather than merely logic. Machine learning, which determines the current wave of artificial intelligence, is built on a data-driven machine experience. The resulting artificial legal intelligence may be far more successful in terms predicting the content of positive law. In this article, I discuss the assumptions of law and the rule of law and confront them with those of computational systems. As a twin paper to my Chorley lecture on *Law as Information*, this should inform the extent to which artificial legal intelligence provides for responsible innovation in legal decision making.

Keywords: artificial legal intelligence, legal services, political economy, machine learning, Rule of Law, speaking law to power, legal theory, speech act theory, information theory, cybernetics, legal protection by design

1 Introduction: does artificial legal intelligence challenge the conceptual foundations of the law?

Artificial legal intelligence has recently been described as a disruption of the legal services market, the legal profession and prevalent business models, though scepticisms regarding overly broad claims and objections against premature replacement of human legal expertise have surfaced.[1] Disruption of the legal profession, however, has been going on for quite some time, based on the commodification of legal services that resulted from the practice of billable hours, combined with the complexity of an increasingly global business practice that finds itself confronted with overlapping jurisdictions including major challenges in terms of dynamic tax law, environmental law and the more.[2] It seems that in the context of a transnational corporation it has become next to impossible to map the multilevel framework of partially contradictory legal requirements that determine corporate liability. At the same time governments are struggling to enact a regulatory system that stays ahead of the myriad of legal constructions developed by corporate law firms to avoid taxation and other costly compliance issues.[3] Also, in

[1] Mainly positive: Richard Susskind, *The End of Lawyers? Rethinking the Nature of Legal Services*, Revised edition (Oxford: Oxford University Press, 2010); Daniel Martin Katz, 'Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry', *Emory Law Journal* 62 (2013); Lawrence B. Solum, 'Artificial Meaning', *Washington Law Review* 89, no. 1 (March 2014): 69–86; Anthony J. Casey and Anthony Niblett, 'Self-Driving Laws', *University of Toronto Law Journal* 66, no. 4 (11 November 2016): 429–42; Albert H. Yoon, 'The Post-Modern Lawyer: Technology and the Democratization of Legal Representation', *University of Toronto Law Journal* 66, no. 4 (11 November 2016): 456–71; Benjamin Alarie, 'The Path of the Law: Towards Legal Singularity', *University of Toronto Law Journal* 66, no. 4 (11 November 2016): 443–55. More or less cautious: Harry Surden, 'Machine Learning and Law', *Washington Law Review* 89, no. 1 (March 2014): 87–115; Brian Sheppard, 'Incomplete Innovation and the Premature Disruption of Legal Services', *Michigan State Law Review* 2015, no. 5 (1 January 2016): 1797; Dana Remus and Frank S. Levy, 'Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law' (Rochester, NY: Social Science Research Network, 27 November 2016).

[2] Yoon, 'The Post-Modern Lawyer', 458-460.

[3] 'It Knows Their Methods. Watson and Financial Regulation', *The Economist*, 22 October 2016,

some jurisdictions access to legal counsel for those not extremely rich has become highly problematic due to the high costs of effective legal advice and representation.[4]

Thinking of law in terms of a market of legal services aligns with an implicit adherence to the basic tenets of law and economics. The latter goes with an understanding of law as just one particular – and maybe even cumbersome – example of governmental regulation that should prove its worth in comparison to other types of regulation that may be more effective and/or efficient.[5] The concomitant 'regulatory paradigm' thinks in terms of influencing the behaviour of people,[6] for which one could also use economic measures, techno-regulation or various types of nudging,[7] which may all prove more effective than legislation or adjudication. Such an economic understanding of law has implied a turning point for the idea of lawyers as dignified stewards of individual justice and societal order. The idea that such stewardship has gone awry was aptly described in Kronman's *The lost lawyer*,[8] though one can criticize its paternalistic and somewhat naïve understanding of the role of lawyers and law firms in the 19th and 20th century.[9] Nevertheless, Kronman and others foresaw or simply documented the turn to commodification, which enables legal scholars to discuss 'the law' as equivalent with 'the legal services market', thus validating the Chicago schools of law and economics (whether based on rational choice and game theory, or on the new school of behavioural economics).

If the perspective is economics, however, I would prefer an analysis of the political economy that drives the 'legal services market', because I am not sure we should take the assumptions underlying the narrative of commodified legal services for granted.[10] Instead of investigating how artificial intelligence will reconfigure the legal services market, I would then ask the preliminary question of how law does and should constitute and regulate such economic markets, while sustaining the checks and balances of the Rule of Law that should define them. My research is focused on how law co-determines the playing field of economic markets and on

---

http://www.economist.com/news/finance-and-economics/21709040-new-banking-rules-baffle-humans-can-machines-do-better-it-knows-their-methods.

[4] Some of the rhetoric around new business models for legal services promises a democratisation of legal knowledge, e.g. Yoon, 'The Post-Modern Lawyer'; Susskind, *The End of Lawyers?* If artificial legal intelligence is going to be part of the trade secret or IP rights of big tech firms this is not a likely scenario.

[5] Still a canonical text: Richard A. Posner, 'The Decline of Law as an Autonomous Discipline: 1962-1987', *Harvard Law Review* 100, no. 4 (1 February 1987): 761–80, doi:10.2307/1341093.

[6] J. Black, 'Critical Reflections on Regulation', *Australian Journal of Legal Philosophy* 27 (2002): 47–226.

[7] Lawrence Lessig, *Code Version 2.0* (New York: Basic Books, 2006); Ronald Leenes, 'Framing Techno-Regulation: An Exploration of State and Non-State Regulation by Technology', *Legisprudence* 5, no. 2 (2011): 143–69; Cass R. Sunstein, *The Ethics of Influence: Government in the Age of Behavioral Science* (Cambridge University Press, 2016).

[8] Anthony T. Kronman, *The Lost Lawyer : Failing Ideals of the Legal Profession*, Reprint edition (Cambridge MA; London: Belknap1995), also: Mary Ann Glendon, *A Nation under Lawyers* (Cambridge: Harvard University Press, 1996); Frank Pasquale, 'Automating the Professions: Utopian Pipe Dream or Dystopian Nightmare?', *Los Angeles Review of Books*, 15 March 2016, https://lareviewofbooks.org/article/automating-the-professions-utopian-pipe-dream-or-dystopian-nightmare/.

[9] E.g. from a feminist perspective Minna Kotkin, 'Professionalism, Gender and the Public Interest: The Advocacy of Protection', *St. Thomas Law Review*, 1 October 1995, 157–74.

[10] Julie E. Cohen, 'The Regulatory State in the Information Age', *Theoretical Inquiries in Law* 17, no. 2 (27 July 2016). Dana Remus, 'Reconstructing Professionalism' (Rochester, NY: Social Science Research Network, 30 April 2015).

Electronic copy available at: https://ssrn.com/abstract=2983045

how it informs the constitutional or constitutive relationship between the public, the social and the private that is the condition of possibility for a durable, fair and thriving market place. Instead of looking at law as a product or service, or as something that regulates a given population, I investigate law as constitutive of societal defaults. Notably those defaults that ground an effective Rule of Law, seen as a system of checks and balances that gives effective standing to those subject to law, affording them effective legal remedies to contest a violation of their fundamental rights.[11] Based on this perspective I will provide a tentative answer to the question of whether artificial legal intelligence challenges the conceptual foundations of the law.

In the next section I dig into the conceptual foundations of the law, acknowledging that these foundations co-determine what Latour has termed the 'mode of existence' of the law. The inquiry builds on my Chorley lecture, which investigated 'law as information in the era of data driven agency'.[12] This entails investigating the concept of information, distinguishing the mathematical theory of information that grounds both computer science and artificial intelligence from the notion of meaningful information that is aligned with human language, text and law-as-we-know-it. Based on theories of human language usage, I discuss the mode of existence of law in terms of speech act theory, highlighting the *performativity* of positive law. In the third section I move on to law as computation, confronting the assumptions of the mathematical theory of information that underpins artificial legal intelligence, highlighting the concept of *performance* as pivotal for machine intelligence, and examine *how machine performance may affect law's performativity*. Finally, I will conclude by suggesting that the concept and practice of 'legal protection by design' may contribute to sustaining law as meaningful information that informs the consequences of our actions, seeking to bring artificial legal intelligence under the Rule of Law.

2 Law as information

2.1 Information, meaning and performativity

According to Raphael Capurro and Birger Hjørland the concept of information comes from the latin *in-formare*, which referred to 'one thing having a formative impact on another thing'.[13] When I say that my life-style informs my health I am not suggesting that my life-style provides information about my health but indicating that my life-style influences or even transforms my health. Precisely because my life-style impacts my health it may also provide information about my health. Capurro and Hjørland trace this second meaning, which refers to 'communicating something about something else', to more recent changes in the use of the notion of information. One could say that information either refers to 'formativeness' or to 'aboutness', while acknowledging their interrelationship. Data that is stored in a computing system without being used can be seen as information-about-something that – until it is employed in some way or another – has no formative impact as such. Nevertheless, it may have already informed the choices people make, or be the result of such choices, and it may still come to inform the algorithms that are trained on the data.

Before moving into the realm of computational systems, we need to check how human beings

[11] Within Europe, the right to an effective remedy against violations of human rights is articulated in art. 13 of the European Convention of Human Rights.

[12] M Hildebrandt, 'Law as Information in the Era of Data-Driven Agency', *The Modern Law Review* 79, no. 1 (2016): 1-33.

[13] Rafael Capurro and Birger Hjørland, 'The Concept of Information', *Annual Review of Information Science and Technology* 37, no. 1 (1 January 2003): 343–411.

share and digest information, as well as how their actions are informed by the information they gather while navigating their world.[14] This 'world' is not just a matter of bricks and bones but consists to a large extent of concepts and institutions that *inform* the consequences of our actions. In the example of life style informing my health one could perhaps replace 'inform' with 'cause', but this is not always possible. Often, the formative aspect of the relationship is based on the performative nature of specific types of language usage. For instance, money is not - or no longer - defined by the value of its physical embodiment (such as with gold coins in an earlier age). Since promissory notes turned into paper money, the concept of money depends on a shared understanding of *what counts as* money. In a sense 'money' is built on thin air.[15] More to the point, it is built on a particular usage of human language. As long as we all agree that the coin, the bill or the numbers in one's electronic bank account *count as* a specific type of currency these coins, bills and electronic bank statements *exist as* money. This is not the result of a postmodern relativism that wrongly suggests that anything and everything is socially constructed (which is of course nonsense),[16] but a consequence of the performative nature of language usage. John Austin and John Searl have highlighted the fact that speech or text does not merely describe but often *does* what it describes.[17] They developed the so-called speech act theory, distinguishing between locutionary speech acts (which contain propositions; 'aboutness') and illocutionary speech acts (which perform what they refer to; 'formativeness'). When a civil servant pronounces the words 'I marry thee (…)' he is not describing what he is doing, but actually doing it. The speech here is an act, it performs the act of marrying a couple and this is just one example of the performative nature of human language use. When I write out a check to another person I am not describing the transfer of money, but actually instructing my bank to provide money to that person upon her request. The piece of paper counts as such an instruction and also counts as a right for that person to receive the money. The check does not merely contain information about that instruction, it transforms the piece of paper into a token that will inform the actions of the bank and the beneficiary. Speech act theory compares well with another inquiry into the performative effects of human language usage, namely semiotics or the study of how signs generate meaning.[18] Semiotics demonstrates that signs never mean anything until they are used by interacting human beings. Language usage implies connecting a system of signs with whatever these signs stand for, refer to or perform. The word 'cup' can be a

---

[14] Navigation is a pivotal aspect of intelligence for mobile beings, as it requires the anticipation of collusion, see e.g. Rodney Brooks, 'Intelligence without Reason', in *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, 1991, 569–95; Rodney Brooks, 'Intelligence without Representation', *Artificial Intelligence* 47 (1991). 'Higher' intelligence derives from navigating our institutional environment.

[15] Pierre Lévy, *Becoming Virtual. Reality in the Digital Age* (New York and London: Plenum Trade, 1998). 68-70. An interesting question is whether bitcoin depends on the use of language in the same way as 'traditional' money, since it seems to entirely depend on self-executing code.

[16] Bruno Latour, *Reassembling the Social : An Introduction to Actor-Network-Theory*, Clarendon Lectures in Management Studies (Oxford ; New York: Oxford University Press, 2005); Bruno Latour, *An Inquiry into Modes of Existence: An Anthropology of the Moderns*, trans. Catherine Porter (Cambridge, Massachusetts: Harvard University Press, 2013).

[17] J.L. Austin, *How to Do Things with Words*, 2nd ed. (Boston: Harvard University Press, 1975); John R. Searle, *Speech Acts, an Essay in the Philosophy of Language* (Cambridge, 1969).

[18] Doede Nauta, *The Meaning of Information* (Mouton, 1972). Semiotics shares some insights with Peircean pragmatism, as Peirce was the founding father of both, see Charles Saunders Peirce, *Selected Writings, Edited with an Introduction and Notes by Philip P. Wiener* (New York: Dover, 1958). Note that the pragmatist maxim defines the expected consequences of the use of a concept as its meaning, highlighting the connection between expectancy theories of language use and its performative effect.

sign that stands for or refers to a physical contraption that allows us to carry liquids, while also enabling us to drink the liquid. To understand the meaning of a cup we need to be aware of other terms, such as liquid, drinking, glass, etc. The word 'education' stands for or refers to a whole network of other terms, but also to buildings, books, teachers and students. Both 'cup' and 'education' depend on shared habits, patterns of behaviour and on the anticipation that practices such as drinking from a cup and enrolling for an education will make sense to others, who live in the same shared world.[19] One of the most interesting features of this shared world, made of bricks and bones, words and habits and fine-tuned mutual expectations, is a persistent discontinuity, an iterant threat of disruption, an abyss that opens up whenever expectations are violated or ignored. Because the meaning of our behaviours, including our words, depend on the web of signs and what they refer to, stand for or 'make up' (as with money), meaning is always shifting. Keeping it stable requires hard work. Meaning is ultimately a matter of joint performance (I cannot create meaning on my own);[20] it is inherently unstable (always depending on how common sense develops) and thereby creates room for productive misunderstandings, different interpretations and reflections on what it is we mean with a word, a concept, a promise, a request or even an order.[21] Therefore, information as aboutness will not do; to sustain our world information must be formative, and where it comes to concepts and institutions it must have illocutionary (performative) force. It is important to note that this particular type of formativeness (constituting meaning) ties in with human language use; we cannot assume that other organisms or data-driven artificial agents are capable of generating meaning in this particular sense.

2.2 Law as information, the performative nature of positive law

2.2.1 The force of law

As should be clear, speech act theory is particularly relevant for the law. The performative aspects of language usage that inform the mode of existence of institutions such as economic markets, marriage or the transfer of property all depend on 'the force of law' or what lawyers call 'legal effect'. The institution of economic markets is enabled by legal precepts such as the freedom to contract and various types of private law liability in case of breach of contract or tort. The legal effect of a marriage is far reaching, notably with regard to creditors and inheritance, and the transfer of property is defined by a whole series of legal conditions that must be met to actually differentiate between 'mere' possession and ownership. From a social science or economic perspective one might believe that the force of law consists of enforcement, by means of court injunctions, finally depending on the power of police (brute force). This, however, is a fatal misunderstanding. The force of law depends on the interlocking of performative speech acts (which are often written, as with legal deeds, legislation and judgments) and the monopoly of violence that looms in the background to ensure that legal effect has the kind of teeth that moral prescriptions lack. If law were only dependent on enforcement it would not be law but administration, discipline or simply a matter of violence.[22]

Legal effect is generated by legal norms, which must be derived from the sources of law,

---

[19] See section 3.3.3.1 'Welt and Umwelt' in my Mireille Hildebrandt, *Smart Technologies and the End(s) of Law. Novel Entanglements of Law and Technology* (Cheltenham: Edward Elgar, 2015).

[20] Ludwig Wittgenstein et al., *Philosophical Investigations*, vol. Rev. 4th (Malden, MA: Wiley-Blackwell, 2009).

[21] Judith Butler, *Giving an Account of Oneself*, vol. 1st (New York: Fordham University Press, 2005).

[22] Obviously, law is in part instituted by violence, see e.g. Robert Cover et al., *Narrative, Violence, and the Law. The Essays of Robert Cover* (Michigan: University of Michigan Press, 1995). Reducing it to violence, however, makes no sense, see e.g. Jacques Derrida, 'Force of Law: The "Mystical Foundation of Authority"', *Cardozo Law Review* 11 (1990): 920–1045.

notably legislation, binding international treaties and case law (the formal sources of law) and legally relevant custom, fundamental legal principles and doctrine (the informal sources of law) that help to interpret the legally binding texts of legislation, treaties and case law. Clearly both types of sources of law interact with and *inform* each other. The sources of law, in that sense, do not merely contain information about the law but institute the law as such. Though *legal knowledge* can be seen as information about the legal effect of one's actions (implying a propositional logic and locutionary speech acts), to understand *legal effect* we must pay attention to how law informs (conditions) the consequences of our actions (implying a learning process and illocutionary speech acts).

2.2.2 Legal certainty, predictability and openness

We can trace a keen awareness of the informative and performative force of the law in Holmes' famous remarks that 'the life of the law has not been logic; it has been experience',[23] defining law as 'the prophecies of what the courts will do in fact and nothing more pretentious'.[24] This links law with experience and prediction, and thus with its role in reducing the uncertainty we may encounter about the consequences of our actions. He may even be seen as a direct precursor of both law and economics and artificial legal intelligence, stating that: 'the rational study of the law the blackletter man may be the man of the present, but the man of the future is the man of statistics and the master of economics'.[25] As Ian Kerr has argued,[26] however, Holmes' argument is about the opposition between logic (inexorable certainty) and statistics (calculated uncertainty), highlighting the crucial importance of providing those subject to law with a means to stabilize expectations, thus enabling people to go about their daily lives. As a practicing lawyer, Holmes knew that legal certainty is never given; it depends on the foreseeability of human action and human interpretation, both of which are fundamentally uncertain. Therefore, it is a matter of statistics rather than logic, of experience and learning rather than framers' intention or literal textual interpretation. This implies that legal certainty is dynamic and requires hard work to be sustained. If this is done well, the uncertainty that grounds legal certainty is not a drawback but an asset, it makes sure that legal certainty does not trump justice and instrumentality,[27] two other aims of the law that are at odds with a rigid interpretation of legal text. Uncertainty is inherent in what Herbert Hart famously described as the open texture of fundamental legal concepts,[28] safeguarding the creative and highly artificial nature of human intercourse as well as legal constructs. As John Dewey has noted, the artificial nature of law must not be equated with its supposedly fictitious nature.[29] An artificial lake is not

[23] Oliver Wendell Holmes, 'The Path of the Law', *Harvard Law Review* 110 (1997): 991–1009. This ties in with the pragmatist maxim, nt 18 and an expectancy theory of legal norms J.F.G. Glastra van Loon, 'Rules and Commands', *Mind* LXVII, no. 268 (1958): 1–9.

[24] Holmes, 'The Path of the Law'.

[25] Ibid.

[26] Ian Kerr, 'Chapter 4: Prediction, Preemption, Presumption: The Path of Law After the Computational Turn', in *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology* (Abingdon, Oxon, [England]; New York: Routledge, Taylor & Francis, 2013), 91–120.

[27] On the tension between these three 'antinomian' goals: Gustav Radbruch, 'Legal Philosphy', in *The Legal Philosophies of Lask, Radbruch and Dabin* (Boston & London: Harvard University Press, 2014), 44–224; Mireille Hildebrandt, 'Radbruch's Rechtsstaat and Schmitt's Legal Order: Legalism, Legality, and the Institution of Law', *Critical Analysis of Law* 2, no. 1 (15 March 2015), http://cal.library.utoronto.ca/index.php/cal/article/view/22514.

[28] H.L.A. Hart, *The Concept of Law* (Oxford: Clarendon Press, 1994).

[29] John Dewey, 'The Historic Background of Corporate Legal Personality', *The Yale Law Journal* 35, no.

an imaginary lake but an act of creation and construction;[30] the discussion should be on how to create and what to construct. This also involves the relationship between performativity and legal certainty.

In an interesting article in a previous issue of this journal,[31] Benjamin Alarie has emphasized the potential of machine learning where it comes to ensuring legal certainty, taking an example from an overly complicated tax law, notably the distinction between an employee and an independent contractor. The distinction is highly relevant for the legal effect generated by tax law, both for corporations and the individual persons that are employed or contracted. The corpus of case law that helps to inform us about the distinction is thus extensive, casuistic and complicated that many billable hours must be spent to advise corporations. Obviously, the space for negotiation on the side of the employee/contractor is not so great and the consequences of misqualification as either an employee or an independent contractor may be hazardous for them also (think of non-declared profits, fines for not keeping administration). According to Alarie, artificial legal intelligence enables a fast and refined prediction of the relevant legal effect, based on a granular analysis of relevant case law by machine learning algorithms. These are trained on the case law by legal experts and data scientists, optimizing the mathematical relationship between input (the corpus of relevant legal texts) and output (the judgements) to a point where the algorithms get it right 98% (which is better than the human lawyers who trained the algorithms). Note that the machines are merely simulating mathematically what human reasoning has come up with in written legal text. Nevertheless, achieving legal certainty as to one's legal obligations under tax law, is a crucial condition for individual freedom, as Jeremy Waldron has insisted:

> There may be no getting away from legal constraint in the circumstances of modern life, but freedom is possible nevertheless if people know in advance how the law will operate and how they have to act if they are to avoid its application. Knowing in advance how the law will operate enables one to make plans and work around its requirements.[32]

However, whereas Alarie seems to believe that a quasi-mathematical accuracy of prediction is the final goal of legal advice,[33] Waldron continues:

> The institutionalized recognition of a distinctive set of norms may be an important feature. But at least as important is what we do in law with the norms that we identify. We don't just obey them or apply the sanctions that they ordain; we argue over them adversarially, we use our sense of what is at stake in their application to license a continual process of argument back and forth, and we engage in elaborate interpretive exercises about what it means to apply them faithfully as a system to the cases that come before us.[34]

Waldron refers to Neil MacCormick as

> pointing out that law is an argumentative discipline and no analytic theory of what law is and what distinguishes legal systems from other systems of governance can afford to ignore this

6 (1926): 655-6 (nt 1).

[30] Ibid.

[31] Alarie, 'The Path of the Law' at 448-451.

[32] Jeremy Waldron, 'The Rule of Law and the Importance of Procedure', *New York University Public Law and Legal Theory Working Papers*, 1 October 2010, http://lsr.nellco.org/nyu_plltwp/234 at 19.

[33] Alarie, 'The Path of the Law'.

[34] Waldron, 'The Rule of Law and the Importance of Procedure' at 20.

> aspect of our legal practice, and the distinctive role it plays in a legal system's treating ordinary citizens with respect as active centers of intelligence.[35]

This is no surprise as MacCormick applied speech act theory to better understand the nature of law and its crucial role in shaping our institutional environment.[36] Law as information must, therefore, be understood as referring to law as a coherent web of speech acts that inform the consequences of our actions, itself informed by the triple tenets of legal certainty, justice and instrumentality that hold together jurisdiction (the force of law), community (even if between strangers) and instrumentality (the policy objectives of the democratic legislator).The antinomian tension between these three tenets require a trained and reflective judgment that is neither mathematically consistent nor based on subjectivist psychology; the judgement is defined by the performative effect is anticipates.

Alarie is betting on what he calls 'legal singularity',[37] the moment that artificial legal intelligence will get it right all the time, thus eradicating any and all legal uncertainty. He seems, however, to mistake the mathematical simulation of legal judgment for legal judgment itself. Whereas machines may become very good in such simulation, judgment itself is predicated on the *contestability* of any specific interpretation of legal certainty in the light of the integrity of the legal system – which goes way beyond a quasi-mathematical consistency.[38] Alarie refers to Rawls' 'reflective equilibrium' as the final touchstone for legal singularity, which assumes (1) that simulation is equivalent with what it simulated and (2) that while legal judgment may still be intractable it is computable.[39] The second assumption ignores the fact that machine learning can be performed in multiple ways, leading to different outcomes. The fact that computability can be 'done' in different ways will 'simply' displace the adversarial debate to the level of training sets, types of algorithms and the design of the hypotheses spaces.[40] On top of that the very fact that computability implies translation and simulation also implies that judgments are not computable if computability is understood as a perfect translation. The first point (mistaking a simulation for what is simulated) turns on the difference between the *performativity of a legal judgment* by a court of law and the *performance rating* of machine learning algorithms. These are two entirely different 'things', requiring a more in-depth inquiry into law as computation.

3 Law as computation

3.1 Computation, feedback and performance

The mathematical theory of information that grounds computer science builds on the second meaning of information (communicating aboutness); it aims to describe how certain discrete bits of content (aboutness) can be transferred from A to B as fast as possible, at low cost,

---

[35] Ibid at 20.

[36] Neil MacCormick, *Institutions of Law: An Essay in Legal Theory* (Oxford University Press, 2007).

[37] This is a reference to Kurzweil's expectation that machine intelligence will supersede human intelligence, creating a hybrid intelligence where human and machine can no longer be distinguished, cf. Ray Kurzweil, *The Singularity Is near : When Humans Transcend Biology* (New York: Viking, 2005).

[38] On the complexities of matching legal certainty with justice and purpose in Radbruch's legal philosophy, see Hildebrandt, 'Radbruch's Rechtsstaat and Schmitt's Legal Order'. On the difference between consistency and integrity Ronald Dworkin, *Law's Empire* (Glasgow: Fontana, 1991).

[39] Something is computable if it can be translated into machine readable data. Due to complexity problems that are computable may nevertheless be intractable when it would take too long to calculate a solution.

[40] Alarie seems to recognize this and foresees a 'legal arms race, with the machines playing a central role in the escalation of tensions' between governments and professional legal advisors working with machine learning sytems, Alarie, 'The Path of the Law', 451.

without changing the content, and without providing access to anyone other than either A or B. The mathematical theory of information was developed by Claude Shannon during the Second World War, with an eye to fast and reliable transfer of secret messages.[41] Shannon was not interested in the content itself but focusing on a type of encryption that would enable to ensure the sameness of the content with as little signals as possible (compression). To achieve the sparest use of signals he defined information in terms of surprise and uncertainty; his goal was to calculate which signals can be left out without distorting the content. In the end this resulted in a mathematical definition of information, based on the concept of entropy.[42] A dataset with an entirely random distribution has complete entropy; the uncertainty is one hundred percent because it is not possible to make any prediction about the distribution. Any deviation of random distribution reduces the entropy and thus reduces uncertainty. To understand the potential impact of data driven applications on the law, we should remember that computer science is built on a separation between signs and their meaning that is radically different from the semantics of human language.

The relationship between information and uncertainty was taken up by Norbert Wiener,[43] the founding father of cybernetics,[44] who was not interested in transferring secret information but in exercising control from a distance. *Cyber* is Greek for 'to steer'; cybernetics is the art or science of remote control over an environment. Wiener focused on how feedback from the environment allows an agent (whether an organism or a machine) to adapt its own behaviour to achieve its goals. The idea is that feedback reduces uncertainty about how the environment may respond to the agent's behaviour; feedback is *information about* the past and the present that supposedly *informs* the future. Herbert Simon, one of the architects of the science of artificial intelligence,[45] explored the crucial role of pattern recognition in feedback. For an organism or a machine to successfully adapt its behaviour it must anticipate how current regularities may play out in the future. Patterns are not the same as regularities; whereas regularities are about the past, patterns aim to reach into the future, they are speculative as they need to move beyond mechanical application.

This brings us to the sub-discipline of machine learning (ML), that is the use of computing systems to detect patterns in data that allow a system to update its own program. The idea is simple: to reach its goals the system must reduce uncertainty about the effects of its own behaviour. To reduce such uncertainty it probes its environment and processes the feedback, reconfiguring its processing algorithms until its goal is reached. Here we return to mathematics, though not to encryption. ML seeks to infer correct mathematical functions to describe patterns in a data set, or between given input data (for instance data on speed or brake behaviour), and given output data, for instance machine readable 'safe driving behaviour' for an autonomous car. The result is a set of – potentially reconfigured - instructions that produces desired behaviour of the system. To develop the right mathematical function data scientists will construct a so-called hypotheses space, consisting of a number of mathematical functions (hypotheses about the relationship between data in a dataset, or between input and output) that

[41] Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1975).
[42] Ibid.
[43] Though they defined information in opposite ways: Shannon equated information with maximum surprise and thus with entropy (random distribution), Wiener equated information with reduction of uncertainty and thus with neg-entropy (structured data). Common sense follows Wiener: information is what reduces uncertainty, the greater the uncertainty the greater the need for information.
[44] Norbert Wiener, *The Human Use Of Human Beings: Cybernetics And Society* (Da Capo Press, 1988).
[45] Herbert A. Simon, *The Sciences of the Artificial* (Cambridge MA: MIT Press, 1996).

are tested, tweaked and again tested, until the systems gets it right most of the time.[46] In his handbook on ML, Tom Mitchell writes:

> a machine learns with respect to a particular task T, performance metric P, and type of experience E, if the system reliably improves its performance P at task T, following experience E.[47]

Much depends on who decide whether the system 'gets it right', which depends on the definition of task T (in machine readable terms) and performance metric P (which determines progress). Obviously, the idea of performance has no relation with that of the illocutionary force or performativity of speech acts. The performance measure is a machine-readable variable that enables the ranking of the machine's behaviour. Similarly, the reliability of ML depends on a number of inevitable trade-offs, notably between volume, relevance, completeness, accuracy and correctness of the training dataset, the dimensionality and aptitude of the hypotheses space, the time taken for iterant testing, and the availability of the relevant domain expertise (how do we know when the machine is 'getting it right'?).

Some have suggested that information is data plus meaning.[48] This highlights the fact that data does not speak for itself; it requires the attribution of meaning by those who use the data, once they use it. It also suggests that information is always meaningful in the above sense. This is clearly not the case. Though many organisms do not speak our type of language, they all depend on information to survive and flourish. Artificial intelligence similarly depends on the processing of information (in the form of digital data points). Neither in the case of organisms nor in the case of intelligent machines does information necessarily imply the attribution of meaning, as it may in the case of humans. Machines work with signs, they do not speak human language; they 'live' in a 'world' made of hardware and software, data and code. Their perception is limited to machine readable data, their cognition is based on computation and manipulation of signs, not on generating meaning (though we may attribute meaning to their output and/or their behaviour). The mode of existence of meaningful information in human society builds on the spoken word, and on handwritten and printed text, each with its own affordances.[49] Meaning differs from information in its dependence on consciousness, self-consciousness and unconsciousness (which depends on consciousness, by definition). Meaning thus depends on the curious entanglement of self-reflection, rational discourse and emotional awareness that hinges on the opacity of our dynamic and largely inaccessible unconscious.[50]

Data, code - and the hyperconnectivity they feed on – do not attribute meaning. Though machine learning has generated feedback mechanisms, these are not similar to human self-

---

[46] On the idea that good enough may be better than best: Herbert Alexander Simon, *Models of Bounded Rationality: Empirically Grounded Economic Reason* (MIT Press, 1982), 295 ff (discussing 'satisficing as often superior to 'optimization'). In the same line Christoph Engel and Gerd Gigerenzer, 'Law and Heuristics: An Interdisciplinary Venture', in *Heuristics and the Law*, ed. Gerd Gigerenzer and Christoph Engel (Cambridge, MA: The MIT Press, 2006), 1–16.

[47] Thomas Mitchell, *Machine Learning*, 1 edition (New York: McGraw-Hill Education, 1997).

[48] Luciano Floridi, *Information: A Very Short Introduction* (Oxford ; New York: Oxford University Press, 2010).

[49] Jack Goody, *The Logic of Writing and the Organization of Society* (Cambridge UK; New York: Cambridge University Press, 1986); Walter Ong, *Orality and Literacy: The Technologizing of the Word* (London/New York: Methuen, 1982); Elisabeth Eisenstein, *The Printing Revolution in Early Modern Europe* (Cambridge New York: Cambridge University Press, 2005).

[50] Butler, *Giving an Account of Oneself*. Siri Hustvedt, *A Woman Looking at Men Looking at Women: Essays on Art, Sex, and the Mind* (Simon & Schuster, 2016).

reflection.[51] Computer systems are now capable of second order – if not n-order – reconfigurations, but they have no self-conscious awareness, they cannot suffer pain or humiliation – even where they excel at simulating such emotions.[52] We are now living with creatures of our own making that can anticipate our behaviours and pre-empt our intent. They inform our actions even if we don't know it and their inner workings seem as opaque as our own unconscious. We have little information about their operations and it is our – largely unconsciously automated[53] – behaviours that inform their often-inscrutable decisions.

Computing systems perform certain functions, based on computation. Their decisions have a normative effect in the sense that they will induce, enforce, inhibit or rule out some of our actions. The crucial issue will be how to differentiate their normative force from the force of law,[54] and how to make sure we do not confuse the performativity of positive law with the performance of artificial legal intelligence.

3.2 Law as Computation: data driven law

3.2.1 Challenging law and Rule of Law

Applying ML to the practice of law involves a plethora of legal services, such as e-disclosure, predictive forensics, assessment of evidence, case law analysis, argumentation mining, analysis of applicable law, and quantitative legal prediction. Much of its success is due to a variety of ML techniques that perform natural language processing (NLP), seeking to develop statistically accurate relationships between an input (documents that are potentially relevant for evidence, case law, legal briefs or memos, doctrinal text, legislation and other types of regulation) and a desired output (relevant documents, relevant lines of argument, precedent, doctrine, applicable legislation or regulation). To the extent that the techniques are used for prediction the output will mostly concern the outcome of upcoming judgements, decisions of public administration, or even the outcome of judgements by particular judges or administrators.[55]

The shift from law as information to law as computation raises a number of issues when it comes to legal intelligence. In light of the concerns outlined above, I will discuss four implications that may disrupt the concept and the Rule of Law: (1) the opacity of ML software may render decisions based on its output inscrutable and thereby incontestable; (2) the shift from meaningful information to computation entails a shift from reason to statistics, and from argumentation to simulation; (3) in the process of developing and testing data driven legal intelligence a set of fundamental rights may be infringed, compromised or even violated, notably the right to privacy, to non-discrimination, to the presumption of innocence and due process, while also impacting consumer and employee protection and competition law. Finally, I argue that, (4) to the extent that the algorithms become highly proficient – due to being trained by excellent domain experts in law – lawyers may outsource part of their work, as a result of

---

[51] Alarie, 'The Path of the Law' at 454 speaks of ALI achieving Rawls' 'reflecting equilibrium'. They can, however, only attempt to simulate this – whether it is reached and how the algorithms should be trained to reach it will always depend on human judgment.

[52] Rosalind Picard, *Affective Computing* (Cambridge, MA: MIT Press, 1997).

[53] Ran R. Hassin, James S. Uleman, and John A. Bargh, *The New Unconscious*, Oxford Series in Social Cognition and Social Neuroscience (New York: Oxford University Press, 2005).

[54] M Hildebrandt, 'The Force of Law and the Force of Technology', in *The Routledge International Handbook of Technology, Crime and Justice*, ed. M.R.P. McGuire and Holt (Routledge, 2017), 579-608.

[55] Katz, 'Quantitative Legal Prediction'; Nikolaos Aletras et al., 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective', *PeerJ Computer Science* 2 (24 October 2016): e93; Daniel Martin Katz, Michael J. Bommarito Ii, and Josh Blackman, 'A General Approach for Predicting the Behavior of the Supreme Court of the United States', *PLOS ONE* 12, no. 4 (12 April 2017): e0174698.

which they may deskill as the software achieves high levels of accuracy. At some point it may be difficult for these lawyers to check whether the software 'gets it right', confronting us with a new Catch22. The latter may be the most foundational implication of artificial legal intelligence.

The opacity of ML operates at different angles.[56] First, the opacity can be due to deliberate hiding of the source code and the relevant training and test data, as they may be part of the trade secret or IP rights of a law firm or the tech firm that developed the software. After having invested in the algorithms, which may have been trained by expensive legal experts, companies may not wish to disclose the inner workings of their systems. Such hiding raises issues with the contestability of the output of the software, rendering difficult or even impossible to test its reliability and to investigate the assumptions that inform the development of the algorithms. Software verification is a subdiscipline of computing science, engaging in either mathematical or empirical verification (preferably both).[57] It can verify mathematical proof of the system by inquiring into the hypotheses space, checking for potential overfitting or overgeneralization, measuring the relevance of the optimization in view of the data used and the given purpose. If no access is given to all this, one could still do empirical testing of the input and the output, irrespective of how the system came to its results. If the software and/or the training data are behind trade secret of IP rights, however, it will not be easy to perform such empirical testing as too much information may be missing. This warrants a legal obligation to work with open source software and to provide access to specified watchdogs capable of testing and contesting the software as well as the training and test data. We must take into account that much of this software will be developed by big tech firms such as IBM and of myriad of startups,[58] not necessarily by practicing lawyers, meaning there is no vested interest in or experience with issues of Rule of Law. Second, the opacity may be due to the fact that neither domain experts in law nor those subject to law have developed any skills to scrutinize these systems. Whereas most of us have learnt to read and write, we were not trained to 'read' and 'write' statistics, we cannot argue against the assumptions that inform ML applications, and we miss the vocabulary that frames 'training sets', 'hypotheses space', 'target functions', 'optimization', 'overfitting' and the more. So, even if experts manage to verify the software, most of us lack the skills to make sense of it; we may in fact be forced to depend on claims made by those who stand to gain from its adoption and on the reputation of those offering this type of data driven legal services. This is also the case when we buy and drive a car, but the reliability here is easier to test. We recognize a car crash and would not appreciate a car that drives us from A to C if we want to get to B; with legal intelligence we may simply not detect incorrect interpretations.

The shift from reason to statistics, and from argumentation to simulation is the most obvious

[56] Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms', *Big Data & Society* 3, no. 1 (1 June 2016); Mireille Hildebrandt, 'Profile Transparency by Design: Re-Enabling Double Contingency', in *Privacy, Due Process and the Computational Turn (*ed. M Hildebrandt and E. De Vries (Abingdon: Routledge, 2013); Bert-Jaap Koops, 'On Decision Transparency, or How to Enhance Data Protection After The Computational Turn', in *Privacy, Due Process and the Computational Turn*, ed. Mireille Hildebrandt and Katja De Vries (Abingdon: Routledge, 2013), 196.
[57] A. Ramanathan et al., 'Integrating Symbolic and Statistical Methods for Testing Intelligent Systems: Applications to Machine Learning and Computer Vision', in *2016 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2016, 786–91.
[58] 'Legal Tech Market Map: 50 Startups Disrupting The Legal Industry', *CB Insights - Blog*, 13 July 2016, https://www.cbinsights.com/blog/legal-tech-market-map-company-list/.

issue for those subject to law and should raise awareness with legal practitioners and scholars. Quantified legal prediction (QLP) can only simulate legal decision-making,[59] it processes signs (data) not meaning (content). The inferences it makes are based on mathematical functions that optimize the mathematical relationship between data points in certain types of legal text (or data points regarding specific judges, courts or administrators) and data points that represent other legal text (for instance a variable that depicts whether an injunction is provided, a violation confirmed or compensation determined). QLP does not argue as lawyers do, it measures and processes elements of their usage of natural language, thus providing feedback on potentially relevant patterns. It feeds on legal argumentation and decision making and does not add its own arguments, other than the numerical rearticulation and mathematical functions that connect the dots. To 'read' the relevant statistics one would have to break the spell of the second type of opacity, as discussed above, and develop the adversarial skills of 'speaking' statistics. This does not necessary imply less of Foucault and Dworkin and more of Shannon and Susskind, as Katz suggests,[60] but should for instance consider that the different schools of statistics generate different outcomes that will make a real difference, especially once nearly anything 'everyware'[61] turns into a data driven Internet of Things. Legal education should make sure that 'reading' statistics is an augmentation of human acuity, and a contribution to an extended human reason, not its replacement. Neither lawyers nor those subject to law should become statisticians, but both should learn to speak the language of variables, functions, correlations, training sets, hypotheses space, while getting the hang of supervised and unsupervised machine learning algorithms. This should ultimately help us *to speak law to the power of statistics*, by which I do not mean to take a silly moralistic approach to ML, but refer to a serious understanding of where the bias inherent in ML applications becomes relevant in terms of the law. For instance, when indirect prohibited discrimination is detected or disproportional processing of personal data that infringes private life, or, when unwarranted monitoring of groups (smart policing, predictive forensics) meddles with the presumption of innocence.[62] This should resolve any tendency to replace legal reasoning with statistical computation, while acknowledging that statistics may contribute to better informed legal reasoning – if done well.

Quasi invisible infringements of human rights and other types of legal protection will be enabled if legal intelligence is taken for granted, notably when it is meant to replace instead of inform legal enaction. We should not fool ourselves, however, by taking for granted that current legal practice entails no such infringements. If legal intelligence is designed and employed to visualize previously invisible discrimination, unsubstantiated conclusions, missing facts, hidden arguments or arbitrary decision-making that flies in the face of both legal certainty and justice as fair and equal treatment, we should applaud it and learn how to 'read' it as such. We should,

---

[59] Katz, 'Quantitative Legal Prediction'.

[60] Daniel Martin Katz, 'The MIT School of Law? A Perspective on Legal Education in the 21st Century', *Univerity of Illinois Law Review*, no. 5 (2014). On the lack of methodological integrity of the use of statistics in social science e.g. https://errorstatistics.com/2016/11/08/gigerenzer-at-the-psa-how-fisher-neyman-pearson-bayes-were-transformed-into-the-null-ritual-comments-and-queries-i/, though I am all for serious introductions to Shannon, Wiener, Gigerenzer and the more. An interesting example of adversarial statistics: https://jasp-stats.org.

[61] Adam Greenfield, *Everyware. The Dawning Age of Ubiquitous Computing* (Berkeley: New Riders, 2006).

[62] Solon Barocas, Sophie Hood, and Malte Ziewitz, 'Governing Algorithms: A Provocation Piece', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 29 March 2013), https://papers.ssrn.com/abstract=2245322; Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact', *California Law Review* 104 (2016): 671–732.

however, not be naïve about the design of legal intelligence. Other than Larry Ribstein, as quoted by Katz,[63] I am not sure that the 'harsh winds of the market' will necessarily blow in the right direction, considering who will be funding the architectures of legal intelligence. If IBM's robolawyer (as their application of legal intelligence has been framed),[64] is trained by top lawyers from major corporate law firms it will reduplicate their concerns and those of their clients. If the contracts between IBM and such law firms give IBM a hold on the data and require law firms to recognize copyright or patent on the software on the side of IBM, we should expect to be confronted by some of the drawbacks that plague other oligarchies driven by large tech companies.

3.2.2 The new Catch22: judgmental atrophy

Both the purported opacity of some types of ML and the schism between mathematical relations and the semantics of natural language imply that the outsourcing of legal intelligence to machines raises the problem of how to evaluate *if and how* these machines are actually 'getting it right'. In his chilling *The Glass Cage*, Nicholas Carr develops a persuasive argument that while smart machines get better in simulating the skills of their human trainers, domain experts may fall prey to gradual deskilling.[65] In a sense, this follows directly from John Searle's famous 'Chinese Room Argument',[66] which demonstrates that while Turing's so-called 'thinking machines' may function as semi-perfect simulators, they never get down to the 'real thing'. They lack consciousness and self-consciousness and even if these were to be epiphenomena of biological evolution,[67] they make 'a difference that makes a difference'[68] *to us* and to *a law that is based on generating both meaning and legal effect*.[69]

In a remarkable article, Federico Cabitza develops a similar argument in relation to medical expertise that is trained into diagnostic ML applications.[70] He narrates how an ML system (Zeb) that is capable of 'reading' radiological images achieves a higher performance than the human radiologists that trained the algorithms, getting it right 85% to 98%. Cabitza raises the question what happens if radiologists routinely consult Zeb, finding that they agree with its diagnosis

---

[63] Katz, 'The MIT School of Law?' at 101.

[64] Jason Koebler, 'Rise of the Robolawyers', *The Atlantic*, April 2017, https://www.theatlantic.com/magazine/archive/2017/04/rise-of-the-robolawyers/517794/.

[65] Nicholas Carr, *The Glass Cage: Automation and Us*, 1 edition (New York: W. W. Norton & Company, 2014).

[66] John Searle, 'Minds, Brains, and Programs', *Behavioral and Brain Sciences* 3, no. 3 (1980): 517–57.

[67] Daniel C. Dennett, *The Intentional Stance*, Reprint edition (Cambridge, Mass.: A Bradford Book, 1989) at 300.

[68] G. Bateson, *Steps to an Ecology of Mind* (New York: Ballantine, 1972) at x.

[69] This is where Solum seems to get carried away, suggesting that machines will produce artificial meaning, just like groups of people do. He fails to register that even if machines could be said 'to produce' such meaning, it is *attributed to the outcome of their operations by* human beings. I would suggest, instead that human beings 'naturally' produce artificial meaning, while machines 'artificially' produce a kind of 'naturalized' meaning (emulating what they infer as the state of the art in legal reasoning). Such 'naturalization' comes close to Alarie's legal singularity, again mistaking a simulation for what is simulated. See Solum, 'Artificial Meaning'.

[70] Federico Cabitza, 'The Unintended Consequences of Chasing Electric Zebras' (IEEE SMC Interdisciplinary Workshop HUML 2016, The Human Use of Machine Learning , 12/16/ 2016, Venice, Italy, 2016), https://www.researchgate.net/publication/311702431_The_Unintended_Consequences_of_Chasing_Electric_Zebras.

most of the time. He suggests:

> Maybe Zeb would help novices learn how to interpret both easy and difficult images; improve the residents' skills more quickly also by challenging them with tricky simulations based on real cases; and teach also expert radiologists how to solve cases that before its arrival would be too difficult for a fast analysis and require several meetings, or worse yet, further examinations for the patient. Conversely, Zeb would perhaps undermine the self-confidence of the expert radiologists (…), after a few times that their interpretation[s] differ from Zeb's and (…) proved to be the correct one. Likewise, it could make the young or less brilliant radiologists more lazy and dependent on its recommendations. Thus the main point here is whether systems like Zeb have the potential to actually deskill or 'spoil' the physicians in the long run.

It seems important to stress that Zeb could actually improve the accuracy of medical diagnosis, while acknowledging that in the long run it could deskill the human specialist. Cabitza enumerates the following unintended consequences in relation to ML-driven decision support systems (MLDSS). He speaks of '*empirical anopsia* and *conversational hypoacousia*, which both regard the representation of the available Experience (E) and the input of the MLDSS'. He describes '*probabilistic tinnitus*, which regards the representation of the outcome of the task (T) and hence the output of the MLDSS', and what he calls '*epistemic sclerosis*, which regards the representation of the performance of the task itself'. Finally he refers to '*metric paroxysm*, that regards the very way performance is assessed (that is, P)' and '*semiotic desensitization, judgmental atrophy* and *oracular rush* which derive from relying too much on the "accurate" output of the MLDSS and its effect on the interpretative work of medical doctors', and lastly he discusses '*metric paroxysm*, that is an ill-grounded trust in numeric measures of the performance of ML-based classifiers', for instance when 'the value of a ML model emerges more from a complex trade-off between its accuracy and its explainability, which must be evaluated qualitatively, rather than by means of a single figure'. Cabitza clearly combines a hilarious depiction of the lack of proper judgment in terms reminiscent of learned medical diagnosis with an inspiring acuity about the drawbacks of human reliance on automation of high level cognitive tasks, adding that an:

> ML approach risk[s] to freeze into the decision model two serious and often neglected biases: selection bias, occurring when training data (the above experience E) are not fully representative of the natural case variety due to sampling and sample size; and classification bias, occurring when the single categories associated by the raters to the training data oversimplify borderline cases (i.e., cases for which the observers do not agree, or could not reach an agreement), or when the raters misclassify the cases (for any reason). In both cases, the decision model would surreptitiously embed biases, mistakes and discriminations and, worse yet, even reiterate and reinforce them on the new cases processed.

What is at stake here is the ability of doctors to recognize when the automation goes awry, misrepresents relevant cases or misinterprets relevant causation. The risk of automating incorrect, imprecise or debatable judgment is not restricted to medical diagnosis. Similar caveats apply when training the algorithms of artificial legal intelligence, where interpretation and subsequent qualification define and determine the legal effect of human intercourse. We need to train our machine to augment human legal intelligence, not to replace it: 'ultimately, this technology is only as good as the training it receives. A poorly trained API [application programming interface, mh] will consistently produce erroneous results, while a well-trained API will reliably generate accurate results'.[71] As law evolves, the software should evolve, and this is where lawyers must come in to make sure fundamental legal principles are taken into

[71] Yoon, 'The Post-Modern Lawyer', footnote 48 at p. 467.

account; otherwise 'self-driving law' may decide on short-cuts that we appreciate in self-driving cars but not in the decision-making infrastructure of human society.

4 Concluding remarks: from 'legal by design' to 'legal protection by design'

In this article I made an attempt to critique some of the pink scenarios of artificial legal intelligence, notably where they seek to emulate a 'complete law' that 'closes the gaps in the law' to achieve 'legal singularity', finally making 'legal uncertainty obsolete'.[72] Though much can be learned from such scenarios, they tell the tales of a totalitarian system, nourishing on a supposedly complete prediction of human behaviour, enabling an approximation of complete compliance, envisaging human behaviour that is 'legal by design'.[73] My argument boils down to the fact that this is not about law, but about discipline or a nefarious type of public administration.

Does this mean that we should forego artificial legal intelligence? As Alarie rightly foresees, current market incentives will probably drive artificial intelligence to a point where the law will be more complex than we can even imagine.[74] Nevertheless, I believe that, as lawyers, we should engage with artificial legal intelligence to make sure it aligns with law and the Rule of Law in a testable and contestable way. As a matter of fact this may save us from ridiculously complex systems outcompeting each other with regard to big business compliance issues.

Bringing artificial legal intelligence under the Rule of Law is not obvious. It will require a specific design of the upcoming computational architecture of our legal systems; indeed, it requires us to reinvent law and the Rule of Law, setting the right defaults, developing the right standards, translating fundamental legal principles into the hardware, the operating systems, the firmware, the software, the applications and the machine learning methodologies we are on the verge of embracing. In my work on privacy, non-discrimination and due process I have coined this 'legal protection by design'. Whereas 'legal by design' is all about compliance, 'legal protection by design' aims to safeguard our ability – as individuals - to challenge automated decision systems, by providing time and space to test and contest the workings of such systems. This will not be an easy task, for the reasons given in section 3.2. We may also ponder the wisdom of catalysing the legal arms race between big business and big regulators, but I believe that there are ways of getting this right, emulating our own legal intelligence in conjunction with ML systems to clarify hidden assumptions, missed arguments, debatable facts and prohibited bias. That way we can speak law to the power of statistics, based on that same power. For those familiar with the paradox of the Rule of Law that is a familiar challenge.

[72] Terms taken from p. 452, the title, 451 and 445 of Alarie, 'The Path of the Law'.
[73] Paul Lippe, Daniel Martin Katz, and Dan Jackson, 'Legal by Design: A New Paradigm for Handling Complexity in Banking Regulation and Elsewhere in Law', *Oregon Law Review* 93, no. 4 (2015), http://papers.ssrn.com/abstract=2539315.
[74] Alarie, 'The Path of the Law' 451.