NOT A SINGLE SINGULARITY

LYRIA BENNETT MOSES

Forthcoming in S Deakin and C Markou (eds), Is Law Computable? Critical Perspectives on Law and Artificial Intelligence (Hart Publishing, 2020)
[2020] *UNSWLRS* 36

UNSW Law
UNSW Sydney NSW 2052 Australia

Not a Single Singularity

Lyria Bennett Moses

Forthcoming Simon Deakin and Christopher Markou (eds) Is Law Computable? Critical Perspectives on Law and Artificial Intelligence (Hart Publishing 2020), https://www.bloomsburyprofessional.com/uk/is-law-computable-9781509937066/.

Director, Allens Hub for Technology, Law and Innovation and Professor, UNSW Law

Analysing affordances and limitations of artificial intelligence (AI) in the context of law is a multi-dimensional puzzle. The most prominent dimension is time, where we travel from descriptions of today's applications, through a relatively clear vision of the near-horizon, to a future sometimes described in terms of singularity – a fully automated legal system where machines are better than lawyers and judges and exponentially improving. Such future imaginaries are difficult to debate – because they lie beyond current knowledge and techniques, any affordance can be hypothesised, while any posited limitation could be overcome. Despite this, scholars have argued that there are limits that cannot be cured by greater processing power, newer AI techniques, or more data.

This chapter does not purport to answer the question 'is law computable?'. Instead, it outlines a means of visualising changes in the computability of aspects of law over time. Rather than visualising improvements in machine intelligence along a single scale towards a single singularity, it describes an evolving solid, expanding around a three-dimensional grid. Changing how we draw the future of AI in law opens up three distinct questions that can be asked at a particular point in time: (1) what is available and to what extent do existing applications replace humans in performing legal tasks and administering law; (2) what are the affordances and limitations of current AI techniques with potential application to law; and (3) in what circumstances should non-human systems be deployed to perform legal tasks or administer the law? These are, respectively, questions in the realms of *is*, *can* and *ought* or, alternatively, availability, capability and legitimacy.

By understanding the problem as three-dimensional, it is easier to see that the legal singularity is not a single point towards which one inevitably climbs. One can progress far along the *x*- and *y*- axes without truly being in a position to automate the role that lawyers, and particularly judges, play in a society governed by the rule of law. Nevertheless, there will be many legal functions that can and will be automated so that the practice of law will change. These mini-upheavals will cause disruption, in courts, in legal practice, and in legal education. But there is no single singularity.

Keywords: Artificial Intelligence, Legal Singularity, Legal Autonomation, Robot Judges, Rules as Code

I. Introduction

Analysing affordances and limitations of artificial intelligence (AI) in the context of law is a multi-dimensional puzzle. The most prominent dimension is time, where we travel from descriptions of today's applications, through relatively clear vision of the near-horizon, to a future sometimes described in terms of singularity – a fully automated legal system where machines are better than lawyers and judges and exponentially improving. Such future imaginaries are difficult to debate – because they lie beyond current knowledge and techniques, any affordance can be hypothesised, while any posited limitation could be overcome. Despite this, scholars have argued that there are limits that cannot be cured by greater processing power, newer AI techniques, or more data.

This chapter does not purport to answer the question 'is law computable?'. Instead, it outlines a means of visualising changes in the computability of aspects of law over time. Rather than visualising improvements in machine intelligence along a single scale towards a single singularity, it describes an evolving solid, expanding around a three-dimensional grid. Changing how we draw the future of AI in law opens up three distinct questions that can be asked at a particular point in time: (1) what is available and to what extent do existing applications replace humans in performing legal tasks and administering law; (2) what are the affordances and limitations of current AI techniques with potential application to law; and (3) in what circumstances should non-human systems be deployed to perform legal tasks or administer the law? These are, respectively, questions in the realms of *is*, *can* and *ought* or, alternatively, availability, capability and legitimacy.

The computability of law can then be plotted as a solid in those three dimensions. Each axis comprises legal tasks otherwise performed by human paralegals, lawyers and judges, and the solid changes over time from invisible (a purely human legal system) to a map of what tasks are being, can be and ought to be performed with current technology. The solid need not always grow. For example, demonstration of limitations or harms of actual or hypothesised techniques, previously unknown, will impact what is understood to be appropriate and, hopefully, what is deployed. The solid is also unlikely to be spherical – things might be done and possibilities created despite their being unsuitable for some contexts. Further, because the axes represent categories of activities rather than a number line, the solid may be disconnected (depending on how the axes are ordered).

This visualisation remains easier to construct for the past than the future. Much has been written on applications of, and limitations of, particular AI techniques such as expert systems and machine learning algorithms (such as random forests and neural networks). But the abstract idea of AI promises infinite potential, up to and beyond the hypothesised singularity. When lawyers imagine natural language processing and machine learning systems replacing judges, the mystique of these ideas often overpowers any practical familiarity with diverse methods and, importantly, their affordances and limitations. The 'intelligence' that machines are mimicking does not function in the same way as human thought and cannot be measured along a one-dimensional scale. Rather than asking about the future existence of a single legal singularity, one is better off asking concrete questions as to how particular applications (real or hypothesised) have and might change the shape of our solid – what is, can and ought to be automated. Like expert systems and machine learning techniques used to predict the future from historical data patterns, newer tools will be useful or even revolutionary in carrying out some tasks otherwise undertaken by humans, while failing miserably at others. The future shape of our solid ('to what extent is law computable?') remains hard to predict, but describing it forces consideration of all aspects of the challenge.

By understanding the problem as three-dimensional, it is easier to see that the legal singularity is not a single point towards which one inevitably climbs. One can progress far along the *x*- and *y*- axes without truly being in a position to automate the role that lawyers, and particularly judges, play in a society governed by the rule of law. Nevertheless, there will be many legal functions that can and will be automated so that the practice of law will change. These mini-upheavals will cause disruption, in courts, in legal practice, and in legal education. But there is no single singularity.

This chapter is structured as follows. Part II discusses what is meant by 'artificial intelligence' and the idea of a legal 'singularity'. Part III *very briefly* describes legal tasks that one might want to automate, including ultimately judging. Part IV introduces the three-dimensional challenge of automation, involving availability, capability and legitimacy. Part V imagines the future shape of legal automation, focusing on the potential for reaching a point where the solid fills the graph, a true 'legal singularity'. Part VI concludes.

II. The idea of artificial 'intelligence' and the singularity

A. Artificial intelligence

Artificial intelligence is a useful term for computer scientists because it describes a kind of problem that might be solved. A person working in artificial intelligence uses a variety of methods (such as neural networks) to perform a species of tasks (those that would otherwise require 'intelligence'). Just as being an expert in contract law implies a kind of shared knowledge, terminology, approach and understanding, so too does being an expert in artificial intelligence.

The use of the term 'artificial intelligence' assumes that we can recognise and measure 'intelligence'. The Oxford Dictionary of English defines intelligence as 'the ability to acquire and apply knowledge and skills.'[1] The Oxford Dictionary of Psychology offers a variety of definitions, starting with 'cognitive ability' then citing scholars for definitions such as 'the ability to carry on abstract thinking' (Terman) and 'the aggregate or global capacity of the individual to act purposively, to think rationally, and to deal effectively with his environment' (Wechsler).[2]

What machines currently do under the heading of 'artificial intelligence' is not always captured by definitions of 'intelligence'. Machines do not 'think', for example, but rather execute their programming. This does not mean that they do not achieve similar functionality at some tasks – I need to 'think' if I want to work out the sum of two numbers, whereas a machine is able to perform the same task more quickly without anything that looks like human thought. Even though machines that 'think' is how some people have conceptualised AI,[3] this remains a futuristic vision.

Many definitions of artificial intelligence thus quite reasonably focus on outcomes rather than internal processes. The classic Turing test for artificial intelligence imagines a human and a machine hidden behind curtains, with a human observer unable to distinguish between them based on the performance of each at a task (such as a text-based conversation).[4] The Oxford Dictionary of Psychology describes artificial intelligence as machines that 'do things normally done by minds'.[5] Similarly, the Dartmouth Summer Research Project on AI in 1955 framed the challenge in terms of doing things that, if a human did them, would require intelligence.[6] Russell and Norvig's *Artificial Intelligence: A Modern Approach*, organises historical definitions into four categories before selecting a series of fields said to comprise AI (natural language processing, knowledge representation, automated reasoning, machine learning, computer vision and robotics), all being techniques that allow an agent to act rationally.[7] The main distinction among these approaches is whether the goal is to emulate humans or the more abstract ambition of rational action, but both focus on observable behaviour rather than internal processes.

While this definition work is useful in defining a field of research and practice, mostly associated with computer science but with interdisciplinary features, it is less helpful as a legal category. It is

---

[1] A Stevenson (ed), *Oxford Dictionary of English*, 3rd online edn, (Oxford, Oxford University Press, 2015) (entry for 'intelligence')

[2] AM Colman, *Oxford Dictionary of Psychology*, 4th edn (Oxford, Oxford University Press, 2015) (entry for 'intelligence')

[3] S Russell and P Norvig, *Artificial Intelligence: A Modern Approach* 3rd edn (Harlow, Pearson, 2014) 3.

[4] AM Turing, 'Computing Machinery and Intelligence' (1950) 59 *Mind* 433.

[5] Above n 2 (entry for 'artificial intelligence').

[6] J McCarthy et al, 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence' (Report, 31 August 1955) http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

[7] S Russell and P Norvig, *Artificial Intelligence: A Modern Approach* (New Jersey, Pearson Education Limited, 3rd ed, 2016) 16–29

neither an appropriate target of technology-specific regulation[8] nor a useful way to understand affordances and limitations of particular tools that are or might be deployed. It is more useful to analyse specific techniques (such as expert systems, machine learning, natural language processing) or even specific methodologies as applied to a given context.

B. The Singularity

An argument has been made that there is a point at which the performance of artificial intelligence will exceed the ability of humans to perform all tasks. This is often narrated in dystopian terms, for example:

> Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an 'intelligence explosion,' and the intelligence of man would be left far behind … Thus the first ultraintelligent machine is the last invention that man need ever make.[9]

Those concerned about the singularity generally turn to the question of whether we can build 'friendly' AI that will act in ways supportive of human flourishing.[10] The challenge, of course, is the difficulty of capturing the complex and contested idea of human value in computer programmed 'intelligence'.[11] There are a variety of organisations focussing the existential risk of artificial general intelligence and how to overcome it, including the University of Cambridge's Centre for the Study of Existential Risk (spawning the Leverhulme Centre for the Future of Intelligence) and the Future of Life Institute. Approaches include negotiating an international treaty mandating prohibition, creating a friendly artificial general intelligence that would prevent the development of other artificial general intelligences, creating an net of artificial general intelligences to police the collective, or augmenting humans to collaborate with or control all artificial general intelligence.[12]

The idea of a singularity suggests a one-dimensional comparison between artificial and human intelligence over time, with a moment in time at which there exists a machine that is more 'intelligent' than humans. At this moment, the machine will surpass humans in all capabilities.[13] In reality, however, we are likely to see machines that are generally intelligent (acting beyond human capability in numerous tasks) before we see machines that outperform humans in *all* tasks. A machine that could outperform Einstein at creative scientific thought would be impressive even if it wrote mediocre poetry. Not even human intelligence is measured on a one-dimensional scale. In the world of education, there is a category of children who are 'gifted but learning disabled,' being '[c]hildren who … exhibit remarkable talents or strengths in some areas and disabling weaknesses in others.'[14] This idea only *seems* paradoxical because we usually imagine intelligence as a single

---

[8] See M Guihot and L Bennett Moses, *Artificial Intelligence, Robots and the Law* (Sydney, LexisNexis, forthcoming 2020) Ch 10.
[9] IJ Good, 'Speculations Concerning the First Ultraintelligent Machine' in F Alt and M Rubinoff (eds), *Advances in Computers*, Vol 6 (New York, Academic Press, 1965) 31.
[10] J Tallinn, 'The Intelligence Stairway' (Sydney Ideas Conference, Sydney, July 2015). https://www.youtube.com/watch?v=BNqQkFg-7AM.
[11] L Muehlhauser, *The Intelligence Explosion* (online, 2011) http://intelligenceexplosion.com/en/2011.
[12] A Turchin, D Denkenberger and BP Green, 'Global Solutions vs Local Solutions for the AI Safety Problem' (2019) 3 *Big Data and Cognitive Computing* 16.
[13] D Wood, Risk Roundup, Risk Group, https://youtu.be/WGtsY8KOa8Q.
[14] S Baum, 'Gifted but Learning Disabled: A Puzzling Paradox' *LD OnLine* <http://www.ldonline.org/article/5939>.

measurement (say, IQ score) rather than as a complex combination of abilities. If AlphaGo (the system that beat human champion Go players) were a human child, it could well be classified as both gifted and learning disabled.

C. Diverse intelligence and milestones

The fact that human intelligence is complex suggests that any notion of a singularity should be multi-dimensional. We are not on a road to a single point in time, but rather meandering through a forest with turns and twists and occasionally getting lost. Software can be awe-inspiring, or it can be buggy and malfunctioning. If we want to understand the role that technology does, can and should play in law, then we need something more than an arrow on a diagram moving unrelentingly upwards.

Clarke has introduced the term 'complementary artefact intelligence' as an alternative to AI.[15] This term recognises that the computers perform well at some tasks (typically those requiring reliability, accuracy and/or speed), and are useful where there are issues of cost, danger or mundanity,[16] but still require human direction. The building of 'smarter' machines is best understood not as a race against humans, but rather as components of human-machine systems that interface effectively, efficiently and adaptably with both humans and other artefacts.[17] Wu makes a similar point in the legal context, arguing that improving machine-human systems is a better focus than human displacement, particularly in the context of the administration of justice.[18]

III. Automation of legal tasks and the legal singularity

Automation of legal tasks is part of a broader conversation around the evolution of work in an age of AI taking place in both scholarship and the broader media.[19] If a machine can be as 'intelligent' as a human (assuming this is a one-dimensional comparison), then it can arguably be as good a lawyer.[20] Alarie has thus introduced the idea of the 'legal singularity' which 'will arrive when the accumulation of massively more data and dramatically improved methods of inference make legal uncertainty obsolete.'[21] Facts, once established, will 'map on to clear legal consequences'.[22] This postulates not only a statement about technology (analogous to the idea of the technological singularity) but also a statement about law (that it is computable, with inputs leading to an output in every case).

The 'legislation as code' or 'machine-readable laws' movement is a vision of *partially* computable law. As an experiment, the Service Innovation Lab in New Zealand trialled the idea of rewriting existing legislation as software code.[23] The goal was to align service delivery and decision making (which is increasingly automated) with the legislation being implemented by writing legislation in a

---

[15] R Clarke, 'Why the World Wants Controls over Artificial Intelligence' (2019) 35(4) *Computer Law & Security Review* 423, 429-430.
[16] Ibid 430.
[17] Ibid.
[18] T Wu, 'Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems' (2002) 119 *Columbia Law Review* 2001.
[19] Eg T Meltzer, Robot Doctors, Online Lawyers and Automated Architects: the Future of the Professions?, *Guardian* (15 June 2014), https://perma.cc/73Q4-WVZA; J Koebler, 'Rise of the Robolawyers' *The Atlantic* (April 2017); JO McGinnis and RG Pearce, 'The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services' (2013) 82 *Fordham Law Review* 3041, 3041.
[20] E Volokh, 'Chief Justice Robots' (2019) 68 *Duke Law Journal* 1135 (2019).
[21] B Alarie, 'The Path of the Law: Toward Legal Singularity' (2016) 66 *University of Toronto Law Journal* 443.
[22] Ibid.
[23] Better Rules for Government Discovery Report (March 2018), https://www.digital.govt.nz/dmsdocument/95-better-rules-for-government-discovery-report/html.

form that software can 'read'.[24] As the New Zealand group determined, rather than taking existing law and re-writing it, it is more effective to change the content of law to that which can be automated.[25] Waddington describes this as 'co-drafting' where 'the legislative drafter is drafting the legislation at the same time as the coder is drafting the coding language' with feedback between them.[26] Such formalisation processes can resolve ambiguities in how sections relate to each other, improving clarity for human as well as machine readers.[27] Designing legislation that can be implemented by a machine has several benefits, including efficiency, transparency (everyone understands how the machine works) and consistency and predictability (everyone can predict the outputs given an input).[28] But it does not work for all kinds of law. In particular, the New Zealand team found:[29]

> *The features of legislation that we identified that are likely to mean that it would be of value to make it available in a machine consumable format are, if the legislation:*
>
> - *involves a calculation*
> - *involves a process that requires factual information to determine application, eligibility, entitlements, or coverage*
> - *prescribes a process that is used repeatedly*
> - *prescribes a compliance process or obligation (for example, regulations that set out 14 different steps that must take place before raw milk can be certified as being fit for human consumption)*
> - *prescribes a process or system that can be delivered digitally.*

This description captures *some* law, but not *all* law or even all legislation. In New Zealand, the case studies used were the *Rates Rebate Act 1973* (NZ) and the *Holidays Act 2003* (NZ), both of which can be imagined as expert systems deploying decision trees and formulas. This example suggests that some laws can be rendered computable based on simple programming techniques.

Using rules as code techniques to render *all* law computable would require changing the content of that law. Waddington's idea of co-drafting can either be cabined (as in New Zealand to particular kinds of law) or expanded across the legislative repertoire. However, the latter would limit the kinds of laws that could be drafted – it would need to remove all discretion, ambiguity and vagueness. While reducing these in some contexts might be desirable, doing so across the board would be a significant limitation. Vague terms can be used to provide flexibility, while contestable terms can ensure evaluation takes place along particular lines, and both allow rule-makers to finesse their disagreement.[30] Discretion allows parliament to ensure that all factors are taken into account in a

---

[24] Ibid.
[25] Ibid.
[26] M Waddington, 'Machine-consumable legislation: A legislative drafter's perspective – human v artificial intelligence' *The Loophole* 21 (July 2019)
[27] SB Lawsky, 'Formalizing the Code' (2017 70 *Tax Law Review* 377, 379.
[28] Ibid.
[29] Better Rules for Government Discovery Report (March 2018), https://www.digital.govt.nz/dmsdocument/95-better-rules-for-government-discovery-report/html.
[30] HLA Hart, 'Jhering's Heaven of Concepts and Modern Analytical Jurisprudence (1970), reprinted in HLA Hart, *Essays in Jurisprudence and Philosophy* 265, 269-70 (Oxford, New York, Clarendon Press, Oxford University Press, 1983); J Waldron, 'Vagueness in Law and Language: Some Philosophical Issues' (1994), 82 *California Law Review* 509, 512-14; JA Grundfest and AC Pritchard, 'Statutes with Multiple Personality Disorders: The Value of Ambiguity in Statutory Design and Interpretation' (2002) 54 *Stanford Law Review* 627.
reprinted in ESSAYS IN JURISPRUDENCE AND PHILOSOPHY 265, 269-70 (1983) ("It is a feature of the human predicament, not only of the legislator but of anyone who attempts to regulate some sphere of conduct by means of general rules, that he labors under one supreme handicap—the impossibility of foreseeing all

wide variety of potentially unanticipated scenarios. The rules as code movement thus accepts that not *all* law can or should be enacted through machine-readable code and that the focus sould be prescriptive rules where there is a use case for automation.[31]

Applying legislation to particular fact scenarios is only one kind of legal task. Various scholars have looked more broadly at the kind of legal tasks that are being or can be automated, in particular Alairie, Niblett and Yoon,[32] Remus and Levy,[33] and Susskind and Susskind.[34] While different authors reach different conclusions about the likely extent of automation over the short, medium and longer term, all agree that there are a variety of tools being used in the delivery of legal services and the performance of legal tasks, and that these will expand over time.

Even technologically unsophisticated systems can help people navigate the legal system. The rules as code example above does not require new AI techniques, only government commitment and proper processes. The 'DoNotPay' chatbot is another example of a simple but effective system – it asks a series of questions then provides advice and documentation to help users avoid traffic and parking tickets.[35] Students in my one-semester course on *Designing Technology Solutions for Access to Justice* have built tools to help teenagers navigate age of consent laws, to help music festival attendees find out the legality of a police search conducted on them, to assist social housing tenants in following the complex process for getting repairs done, and to help human rights workers in the region navigate the maze of relevant international treaties. Expert systems can also be used to collect instructions, complete forms and personalise documents.[36]

Beyond expert systems and other pre-programmed logics lie the possibilities opened up by machine learning and natural language processing, the two branches of AI with significant implications for law. Search engines that locate doctrinally relevant material increasingly rely on these techniques. For example, Ross (associated with IBM Watson) can understand legal questions written in regular English and provide answers based on its knowledge database.[37] In the future, advanced versions of these tools will make legal research significantly more efficient. Machine learning prediction engines, such as Lex Machina and Premonition, search for trends that facilitate prediction of likelihood of success (for different judges), quantification of damages, and likely cost and time of proceedings.[38] Tools that partially automate document review (whether for due diligence or discovery) are also increasingly popular.[39]

---

possible combinations of circumstances that the future may bring. . . . This means that all legal rules and concepts are 'open'; and when an unenvisaged case arises we must make a fresh choice, and in doing so elaborate our legal concepts, adapting them to socially desirable ends.")

[31] See eg digtal.nsw (NSW Government), 'Emerging Technology Guide: Rules as Code', https://www.digital.nsw.gov.au/digital-transformation/policy-lab/rules-code.

[32] BH Alarie, A Niblett and AH Yoon, 'How Artificial Intelligence Will Affect the Practice of Law' (2018) 68 (supplement 1) *University of Toronto Law Journal* 106, 109ff.

[33] D Remus and F Levy, 'Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law' (2017) 30(3) *Georgetown Journal of Legal Ethics* 501.

[34] R Susskind and D Susskind, *The Future of the Professions: How Technology will Transform the Work of Human Experts* (Oxford, Oxford University Press, 2015).

[35] Available at https://donotpay.com.

[36] J Bennett et al, 'Current State of Automated Legal Advice Tools' (Discussion Paper No 1, Networked Society Institute, University of Melbourne, April 2018) 22–25.

[37] See Ross Intelligence, https://rossintelligence.com/.

[38] See https://lexmachina.com and https://premonition.ai.

[39] T Davey and M Legg, 'Machine Learning Disrupts Discovery' [2017] (32) *Law Society Journal (NSW)* 82.

Collectively, these examples demonstrate that some legal tasks are being delegated to machines and that possibilities expand as new AI techniques are introduced. There have, however, been strong critiques of the use of particular tools in particular contexts. For example, the COMPAS tool, deployed extensively in the United States, provides a risk assessment score to judges as a measure of the likelihood that a given defendant will re-offend for use in sentencing.[40] This tool not only fails to operate fairly by some measures (the false positive rate had a disparate impact on African Americans[41]) but it explicitly uses variables that would not ordinarily be considered relevant to sentencing.[42]

The map of where we are thus includes both excellence and failure, both usefulness and harm, both affordances and limitations. It seems less a march towards an inevitable singularity that a journey with choices – to develop AI tools and implement them in an expanding range of contexts. Looking back, we can recognise mirages, such as uses that seemed efficient and fair but were in fact problematic and illegitimate. Understanding our travels to date and where we might go in the future requires a more complex map than a single path upwards towards a single singularity. As Clarke has argued in a broader context and Wu in the legal context, it also requires a break from the idea of humans versus machines to the construction of systems that deploy integrated human and machine components in ways that take advantage of the strengths of each.[43] Moving forwards then is not about approaching the legal singularity as a destination, but rather enhancing the legal system (including its tools and processes) for the benefit of society.

IV. A three-dimensional challenge

If one lists or maps the technologies that have been developed for or deployed in legal contexts, particularly in recent times, it could be narrated as the first steps along an inevitable climb towards a legal singularity. One might think that it is impossible for technological forecasts to be too optimistic. However, when mapping techniques along three separate dimensions – availability, capability and legitimacy – it becomes clear that developments are mostly about volume and technical range, ignoring the more fundamental issues around legitimacy.

A. The *x* axis: Availability of useful tools

The *x*-axis, representing the volume of tools available to perform legal tasks, has moved quickly, although there is a considerable way to go. Most legal technology conferences include a presentation with a slide that sets out all of the legal technology companies operating (either generally or in the relevant jurisdiction) and that slide feels more crowded each year.[44] Tools include those that target marketing to potential clients, analyse and classify documents, analyse data to make predictions about legal costs or changes of success, draft legal documents, answer legal questions, assist with legal research, help with practice and document management and facilitate

[40] See description in *State Wisconsin v Loomis,* 881 NW2d 749 (Wis 2016). The United States Supreme Court denied certiorari on 26 June 2017.
[41] J Angwin et al, 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks' *ProPublica* (online, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
[42] M Zalnieriute, L Bennett Moses and G Willliams, 'The Rule of Law and Automation of Government Decision-Making' (2019) 82(3) *MLR* 425, 447-8.
[43] Clarke, above n 15.
[44] For example, see LawGeex, *Legal Tech Buyers Guide 2019*, https://ltbg2019.lawgeex.com/products-by-category/

collaboration. Law firms are also building products in house or in partnership with technology companies.[45] It is a crowded, but exciting and fast-moving space.

Beyond corporate endeavours, are projects with more communal goals, such as Austlii's DataLex. Austlii (austlii.edu.au) provides free legal information for Australia and, through partnerships, beyond. The original DataLex project ran from 1984 to 2001, creating a platform for developing legal expert systems, incorporating a case-based inferencing mechanism, a full text retrieval system and a hypertext engine.[46] Enhanced DataLex tools are currently available on the AustLII Communities platform, a closed wiki, with developers able to play with building expert systems that can navigate complicated legislation.[47] With further development work, DataLex could be used to convert all Australian legislation into an expert system, with users answering questions (which may require interpretation of terms or legal guidance) to find out how particular laws apply to their personal circumstances. The use of DataLex software is free for students and legal community centres and there is no cost to end-users.[48] The real benefit is that development work can be conducted by lawyers with no coding expertise. However, to date, only a small part of Australia's statutory corpus has been re-written in the format required by DataLex.

Although the *x*-dimension appears straightforward, there is nothing 'easy' about creating legal tools that can replace or assist humans in performing legal tasks. Time can be spent designing, building and marketing legal tools without any expansion along the *x*-axis. There are many examples of poor design and implementation. Consider an interactive mobile phone application built by Victoria Legal Aid offering 'targeted, relevant and free information to young people on legal issues that affected them.'[49] It focused on legal information about consent to sexual intercourse, sexting, and cyberbullying.[50] For example, users could enter personal information to find out whether they were above the age of consent. The application was launched in November 2013 but encountered low install rates and high uninstall rates and was discontinued in November 2015.[51] At that time, only 1095 users had installed the app and only 40 of those had created accounts. Further, the app could no longer run on newer Android phones due to an operating system upgrade. In the evaluation, there were problems with lack of a comprehensive market strategy, failure to consider the value proposition for the client, failure to consider challenges with building on one platform only (Android).[52] Despite its ambitions, the application did not significantly expand the circumstances in which users could obtain automated legal advice, and thus did not give rise to growth along the *x*-axis. This example highlights that what is often at stake in automation is not how technically sophisticated the methodology, but: how buggy the software, how well-designed the tool, processes of implementation and marketing, and building solutions that align with problems not solutions in isolation.

A jurisdiction travels along the *x*-axis though the conception, design and building of useable systems to automate legal tasks. This will draw on existing AI capabilities, requiring also time, quality design and effective implementation. It is crucial that such tools be developed with an understanding of the

---

[45] L Bennett Moses, 'The Need for Lawyers', in K Lindgren, F Kunc and M Coper M (eds), *The Future of Australian Legal Education* (Australia, Lawbook Company, 2018) 355.
[46] G Greenleaf, A Mowbray and P Chung, 'The DataLex Project: History and Bibliography' (3 January 2018), [2018] UNSWLRS 4.
[47] Ibid.
[48] DataLex, Legal inferencing systems (brochure)
[49] Victoria Legal Aid, *Case study of the BELOW THE BELT PHONE APP* (May 2016).
[50] Ibid.
[51] Ibid.
[52] Ibid.

broader socio-technical networks in which they will be deployed, including the needs of potential users.

B. The *y* axis: Evolving capabilities

Each 'type' of AI has both affordances and limitations. As new AI techniques are introduced, automated tools can perform new tasks in the legal domain. The current boom in legal technology is largely related to developments in machine learning and natural language processing, which have facilitated prediction of legal outcomes, automation of document review, and more intuitive legal search products.

The limitations of expert systems were identified decades ago.[53] Expert systems can manage legal domains that are simple (like age of consent laws) and complicated (like tax law), but not those that are complex and unpredictable (such as evaluations of reasonableness).[54] Similar points apply to the rendition of rules as code. In some cases, drafting law in this way would be an improvement. In particular, it reduces the need for those seeking to operationalise 'compliance by design' to each independent (and possibly inaccurate) conversion of laws from text to machine-readable code. Instead, everyone can rely on the government-endorsed machine-readable version of rules. However, 'rules as code' does not purport to deal with discretionary components of legislation; discretion remains outside the system but, once a decision is made, can be looped back into the process. Discretion can only be delegated to a machine if one is satisfied with pure chance[55] or prediction of outcomes based on machine learning trained on historic exercises of discretion or other events. Further, 'rules as code' requires avoidance of language that permits ambiguity and flexibility in favour of programmable logic.[56] As noted above, this is undesirable in some circumstances and is outside the remit of the rules as code movement.

Machine learning can surmount the limitations of pre-programmed rules. In a learning process, one does not need to know a rule. Instead (in supervised learning) one can have historical classifications (such as whether the plaintiff won or lost) and the facts of those cases. The machine can learn to identify the variables (facts) that correspond with the classification (win or lose). Machine learning can be used to 'learn' the weights given in practice to a range of factors known to be relevant. Further, it can learn to identify more complex patterns, as where the presence of one or more factors changes the weight given to a different factor.

The closest machine learning has come to replacing judges is predicting judicial behaviour. Aletras et al have built a tool to analyse textual content from European Court of Human Rights judgments in order to predict the final outcome, a task accomplished with 79% average accuracy.[57] There are flaws in that particular study, in particular the fact that the training set was descriptions of the facts

---

[53] See generally P Leith, 'The Rise and Fall of the Legal Expert System' (2010) 1(1) *European Journal of Law and Technology*, http://ejlt.org//article/view/14/1; R Stamper, 'Expert Systems – Lawyers Beware!' in Stuart S Nagel (ed), *Law, Decision-Making, and Microcomputers: Cross-National Perspectives* (Westport CT, Quorum Books, 1991) 19, 20.

[54] DJ Snowden and ME Boone, 'A Leader's Framework for Decision Making' (2007) *Harvard Business Review* 69 (November issue).

[55] A D'Amato, 'Can/Should Computers Replace Judges?' (1977) 11 *Georgia Law Review* 1277, 1279.

[56] F Pasquale, 'A Rule of Persons, Not Machines: The Limits of Legal Automation' (2019) 87 *George Washington Law Review* 1, 3.

[57] N Aletras et al, 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective' [2016] (2) *Peer Journal of Computer Science* 92.

in the judgments themselves.[58] Any tool that claims it can predict judicial decisions needs to be carefully evaluated, and accuracy rates need to be measured as against pure chance (50%) not pure error (0%).

Machine learning also has limitations, different to those of expert systems. A machine learning algorithm will only predict the outcomes of new cases where they are sufficiently similar to those on which it was trained. Human decision-makers are better at dealing with unanticipated circumstances and using 'common sense' to assess the importance of a new situation. While a computer will only know speeding might be permitted in a medical emergency if it has previously seen this scenario, humans can deduce that this situation may need to be separately assessed even if it is the first time it is encountered.[59] Machine learning is inherently conservative, better at making predictions about the legal system as it currently exists rather than making suggestions for how it should evolve.[60] At most, it can learn a trend that is already present in the data (such as an increase in damages payouts) and project that trend into the future. But it cannot change course – historical legal revolutions such as recognition of tort liability for negligence[61] or the recognition of native title in Australia[62] would not be dreamt up by a system that replicates historic patterns and trends. Some might think this is a good thing (confining judicial innovation) but it is not a mirror for current judicial practice. Further, machine learning is often criticised for not justifying outputs in legally relevant terms.[63]

The question is not only about what machines can (and cannot) accomplish independently. Automated systems often work alongside humans in performing legal tasks, providing legal advice and making legal decisions. Expert systems are often designed to be used by lawyers (who can interpret legal language in particular factual contexts) rather than clients. Machine learning is most commonly used in tasks that would otherwise be assigned to junior lawyers. Machine learning systems must be trained by junior lawyers and remains under the supervision of senior lawyers. It is then used in e-discovery, document classification and clustering, semantic legal search, data-driven document creation, narrative generators, predictive analytics and relevance, negotiation and risk optimisers.[64] These are impressive, and have led to questions about how junior lawyers will gain basic skills without these standard large-scale tasks.[65] But accomplishing more strategic tasks independently will likely require another technological revolution. In the meantime, affordances and

---

[58] F Pasquale and G Cashwell, 'Prediction, Persuasion, and the Jurisprudence of Behaviourism' (2018) 68 (Supplement 1) *University of Toronto Law Review* 63, 68-72.

[59] G Marcus and E Davis, *Rebooting AI: Building Artificial Intelligence we can Trust* (New York, Pantheon, 2019).

[60] D Remus and F Levy, 'Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law' (2017) 30(3) *Georgetown Journal of Legal Ethics* 501, 549.

[61] *Donoghue v Stevenson* [1932] UKHL 100, [1932] AC 562.

[62] *Mabo v Queensland (No 2)* (1992) 175 CLR 1.

[63] F Pasquale, 'Toward a Fourth Law of Robotics: Preserving Attribution, Responsibility, and Explainability in an Algorithmic Society' (2017) 78(5) *Ohio State Law Journal* 5; L Bennett Moses and J Chan, 'Using Big Data for Legal and Law Enforcement Decisions: Testing the New Tools' (2014) 37(2) *University of New South Wales Law Journal* 643.

[64] JO McGinnis and RG Pearce, 'The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services' (2013) 82 *Fordham Law Review* 3041; D Ben-Ari et al, '"Danger, Will Robinson"? Artificial Intelligence in the Practice of Law: An Analysis and Proof of Concept Experiment' (2017) 23 *Richmond Journal of Law and Technology* 2, 31-5; D Remus and F Levy, 'Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law' (2017) 30(3) *Georgetown Journal of Legal Ethics* 501; H Surden, 'Machine Learning and Law' (2014) 89 *Washington Law Review* 87.

[65] McGinnis and Pearce, above n 64, 3065-6. However, one might equally wonder how juniors could manage without the opportunity to retype or manually edit correspondence or deliver paper document folders around the city –tasks that would likely have been common training ground in the past.

limitations need to be assessed for tools operating independently as well as in the more common context of machine-human systems.

C. The *z* axis: Legitimacy and appropriateness of deployment

There are important questions to ask beyond affordances and limitations of particular systems and tools. The existence of a *z* axis is an assertion that not every tool that *can* do something *should* be used in all circumstances. For example, there may be reasons why, even if a system could predict the decision of a given human judge with a high degree of accuracy, we would want the decision to be made by the judge rather than the system.

Some have argued that the existence of a z-axis ultimately comes down to techno-scepticism or sentimentality, and that society should focus on outputs rather than the means through which those outputs are produced. According to this argument, if a system can produce outputs that are assessed as at least as good as those of human professionals, then there is no reason to prefer humans. Volokh argues, for example, that we should assess hypothetical lawyer robots by whether the outputs provide us with what we need, which is persuasiveness at least equal to the average human performing the same task.[66] According to him, it is irrelevant whether a judge reaches decisions by a similar process to a human (applying legal rules to facts), provided that the output is equivalently or more persuasive (according to a panel of evaluators).[67] If Volokh is right, there is no *z* axis, no questions to be asked beyond technical capability (y axis) and practical implementation and deployment (x axis).

To understand the flaw in Volokh's argument, it is worth looking back to earlier thoughts about the limits of automated systems. In 1976, Weizenbaum (author of the language analysis program DOCTOR that could substitute for a psychotherapist), wrote that 'since we do not now have any ways of making computers wise, we ought not now to give computers tasks that demand wisdom.'[68] While we can build machines that learn, we have not yet built one that could be described as wise, albeit that it may be able to predict average behaviour of people who deserve that accolade. Dreyfus had slightly different concerns, about modelling the indeterminacy of the problems addressed by humans.[69] There are a variety of modern arguments against machines generally or machine learning specifically that deploy a range of concepts, but ultimately align with Weizenbaum's concern about wisdom.

Hildebrandt is concerned that those focusing on outputs seem to 'mistake the mathematical simulation of legal judgment for legal judgment itself.'[70] She links judgment with recognition of the contestability of legal interpretation, citing Waldron that we do not just obey laws, we argue about them.[71] According to Hildebrandt, the output of a machine learning system is less practically contestable (due to its opacity) and is not contestable on the same terms – one can argue the statistics but not the reason. Contestability is also reduced within the broader system given the deskilling of those who could conduct a full evaluation in the future. At least in the context of legal

[66] Volokh, above n 20.
[67] Ibid, 1162.
[68] J Weizenbaum, *Computer Power and Human Reason: From Judgment to Calculation* 227 (1976).
[69] HL Dreyfus, What Computers Can't Do (Harper & Row, 1972) 194
[70] M Hildebrandt, 'Law *as* computation in the era of artificial legal intelligence' (2018) 68 (Supplement 1) *University of Toronto Law Journal* 12.
[71] J Waldron, 'The Rule of Law and the Importance of Procedure' in JE Fleming (ed), *Getting to the Rule of Law, NOMOS L* (New York and London, New York University Press, 2011).

decision-making that is integrated into the state, a contestability requirement would render illegimate judgments created through automation without human oversight.

Kerr and Mathen[72] also raise issues that go beyond the alignment of outputs, recognising the importance of the state of mind of actors in the system. They give the example of the oath sworn by federal US judges:

> I, A. B., do solemnly swear or affirm, that I will administer justice without respect to persons, and do equal right to the poor and to the rich, and that I will faithfully and impartially discharge and perform all the duties incumbent on me as, according to the best of my abilities and understanding, agreeably to the constitution, and laws of the United States. So help me God.

This is not a promise not to reach a particular outcome, but a promise to have a particular orientation in the course of making a decision. We do not want a judge to merely *persuade an audience* that they are meeting this standard, we want them to actually take their obligation to heart. As they note 'the mere fact that a machine demonstrates rule-following behaviour does not make it a rule follower'.[73] Such machines cannot truly adopt what HLA Hart described as an internal point of view.[74] Kerr and Mathen conclude 'Legal reasoning also requires being a member of the community, understanding its history, its moral convictions, having a point of view about its current character and having a stake in its future'.[75] Neither expert systems nor machine learning can give rise to systems having such a psychological perspective on their task.

The difference between simulation or prediction of human judges and actual human judges can also be understood by comparing judicial decisions to the conduct of elections. Suppose that pollsters could predict elections with a high degree of accuracy (not the case currently!). Imagine that sampling and statistical methods improve to bring the accuracy to 99%. Then suppose it is argued that elections are expensive to hold and that, for the sake of 1% accuracy, it is better to simply take the pollster's prediction as the outcome. In my view, a country that accepted that logic would no longer be democratic. It is a requirement that the vote be held, not merely simulated or predicted. The point that Hildebrandt as well as Kerr and Mathen are making is similar for judicial decisions – irrespective of the alignment of outcomes, there needs to be a decision-making process that meets particular criteria. For Hildebrandt, this is contestability while, for Kerr and Mathen, it is the ability to psychologically embrace the internal point of view.

These kinds of challenges to automation of legal decision-making are fundamental – they go beyond the limitations of current techniques and pose broader questions about the enterprise itself. It is, however, worth noting that most legitimacy challenges are concrete and relate to specific use-cases with current techniques. Machine learning, in particular, has come under extensive critique, particularly with respect to its alignment with rule of law values.[76] As Surden points out, many of these concerns are contingent.[77] While there are examples of applications that fail to ensure equal treatment under the law (such as COMPAS), there is also the potential for AI to detect bias in historical data and conduct learning within prescribed fairness constraints.

---

[72] I Kerr and C Mathen, 'Chief Justice John Roberts is a Robot' (draft).
[73] Ibid 24.
[74] HLA Hart, *The Concept of Law*.
[75] Above n 72, 39-40.
[76] Zalnieriute et al, above n 42.
[77] H Surden, 'The Ethics of Artificial Intelligence in Law: Basic Questions' *Oxford Handbook of Ethics of AI* (draft chapter).

Pasquale and Cashwell's critique[78] of Aletras et al's prediction of the European Court of Human Rights is arguably contingent in this sense. As they write: 'it is a foundational principle of both administrative and evidence law that irrelevant factors should not factor into a decision'.[79] An automated system looking for correlations in language used in statements of facts and the ultimate decision will potentially take into account facts that ought to be dismissed as irrelevant, likely due to spurious correlation in an inevitably small data set.[80] Even if the system is only used for triage, there are real concerns about its deployment in contexts that affect potential litigants.[81] The rule of law not only cares about outcome, but the mode through which it is generated – the internal logic (human or machine) thus *matters* so that using irrelevant considerations renders the decision illegitimate even if the result is likely to be the same (as noted in high accuracy probabilities). But this is not about the psychology of the decision-making entity. Pasquale and Cashwell are focusing on one specific machine learning algorithm so that it is less clear whether they would accept an algorithm where only weightings and not variables are 'learnt'. An example would be something similar to the New Zealand Risk of Reconviction algorithm which, unlike COMPAS, draws on only specific published variables associated with criminal activity and past interactions with the criminal justice system.[82]

Another critique of automation where there is contingency is alignment with legal ethics. Remus and Levy point out that automated advice systems will generally lack the ethical, law-abiding orientation of human lawyers who give advice. They use the example of tax advice, and the willingness of (some) human lawyers to advise clients to act legally and in line with the spirit of the law even where, statistically, they may be unlikely to face consequences.[83] Other than the interpretation question (what is 'the spirit of the law'?), this could arguably be met by programming. One might well consider 'ethical behaviour' as something that can be assessed purely on the basis of outputs, that is the content of the advice given. In that light, it can be met by (some) systems.

Some objections about legitimacy are broad (applying to all technology in many applications) while others are narrow or contingent (for example, those relating to the legitimacy of particular systems used in particular contexts). Overcoming all types of objections is required to move along the *z*-axis. If we take seriously the concerns of Hildebrandt, Kerr and Mathen, in particular, then we have not travelled very far along the *z*-axis to date. In particular, neither expert systems nor machine learning *ought to* replace human judges in most contexts.

Some argue that human judges ought to be replaced in at least some circumstances. Some have suggested, and Estonia is proposing to adopt, a model where small claims are resolved by AI tools.[84] However, it ought not be simply a matter of the 'importance' of the matter along some objective scale such as total value at stake (in Estonia, the amount is 7000 euro) or classification of the issue as procedural.[85] Smallness ought not to remove a person's entitlement to legitimate justice if that is what they wish to pursue. It may be acceptable where the parties agree to an automated process.

---

[78] Frank Pasquale and Glyn Cashwell, 'Prediction, Persuasion, and the Jurisprudence of Behaviourism' (2018) 68 (Supplement 1) *University of Toronto Law Review* 63.
[79] Ibid 76.
[80] Ibid 76-77.
[81] Ibid.
[82] NZ Government, *Algorithm Assessment Report* (October 2018) 21. The exceptions are age and sex.
[83] D Remus and F Levy, 'Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law' (2017) 30(3) *Georgetown Journal of Legal Ethics* 501, 552-4.
[84] Eric Niiler, 'Can AI Be a Fair Judge in Court? Estonia Thinks So' *Wired* 25 March 2019, https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/.
[85] D'Amato, above n 55, 1289.

After all, most jurisdictions recognise the rights of parties to agree to move outside the state-sponsored justice system, for example by agreeing to arbitration or, for that matter, appearing before Judge Judy or abiding by the result of a coin toss. Provided the contract that delegates decision-making authority to an algorithm is enforceable (including all the jurisdiction-specific matters to be considered), then the delegation should be seen as equivalently legitimate to arbitration clauses. But parties ought not to be forced to accept a simulation of justice without informed consent, even where the total value at stake is relatively small.

Despite measurable progression upon the *x* axis and constant expansion of what is technically possible (represented by the *y*-axis), current techniques such as machine learning and expert systems operating alone fare poorly on the *z* axis, particularly in the context of the administration of justice. There are some objections that can be met with current technologies implemented well, but many which cannot. A judge does not merely find legislation and cases, locate relevant provisions or ratio decidendi, and apply logic to deduce a result. Nor do judges reason statistically from historical data points. Instead, judges are a lynchpin for the rule of law, both in terms of their own decision-making processes and, through building trust and respect, encouraging a broader rule of law culture. Ultimately, judges decide what a decision should be, which is quite different to predicting what the decision will be. Even if machine learning could improve performance against rule of law values such as equal treatment, it could not be contestable on the same terms or take the internal point of view. The strongest progress on the z-axis has been for technologies (such as search engines) that improve human performance.

V. What kind of technology *could* replace judges?

What has been generated so far is a three dimensional solid, growing over time up to the present. This solid has grown most rapidly along the *x*-axis with the large number of businesses expanding into the legal technology space alongside social good projects such as DataLex and rules as code. The growth along the *y*-axis has been sporadic but significant. The main bursts have been associated with expert systems in the 1980s and 1990s followed by machine learning more recently. The solid, is however, rather shallow, with the *z*-axis only embracing voluntary submission to automation, efficiencies in some aspects of legal service delivery, and as a component in human-machine systems (as where it is used to locate relevant materials and improve legal prediction).

The question, then, is whether focusing on a different part of the *y*-axis might enable growth along the *z*-axis for fully autonomous systems. In other words, is there an AI technology that can answer contemporary the concerns about legitimacy? The difference between human judges and machines raised above are not necessarily biology, but rather concern that particular programming techniques do not match the requirements of the rule of law or the psychological attitudinal requirements of judging. It is thus worth asking whether we can progress along the *z*-axis by pursuing different approaches to artificial intelligence.

There are a variety of, at this stage, purely hypothetical sub-fields of AI. In his book on *Superintelligence*, Bostrom proposes a variety of paths to the singularity. One might, for example, be able to take a human judge and perfectly replicate their brain within computer hardware or a computer model. Would such a replicant be in a position to answer the critiques above? Assuming it reasoned the same way as the replicated human judge, it would treat legal propositions as contestable so as to answer Hildebrandt's point. It would also feel bound by rules and take Hart's internal perspective on the law, so as to satisfy Kerr and Mathen. It would not rely on machine learning as a technique but would rather simulate precisely the reasoning of the replicated judge, thus avoiding the issues raised concerning that approach.

There are, however, some concerns that would still apply. Brennan-Marquez and Henderson argue that the situation of the decision-maker is itself important. In particular, they suggest that 'those tasked with making decisions should be susceptible, reciprocally, to the impact of decisions.'[86] Unless machines can be placed in this state *psychologically*, they are arguably not suitable judges. For this to be the case, the replicated brain would need to be placed in a robot that could truly experience the confinement of jail. Even there, one might wonder whether the psychological state of that robot would be sufficiently human that one can describe their feelings about this as truly mirroring those of a defendant in the criminal justice system.

Another matter worth considering is the importance of generational renewal within the justice system. There are ideas currently viewed as central to the rule of law, such as equal treatment of women and racial minorities, that would have felt quite foreign to those born a thousand years ago. Similarly, Oliver Wendell Holmes may have been a famous judge of his time, but 'three generations of imbeciles is enough' would hardly constitute good legal reasoning today.[87] If we uploaded today's judges into artificial brains inside computer systems or robotic bodies, it is unlikely that they would adapt to radical new ideas. Judicial decision-making would become increasingly conservative, possibly leading to stagnation of society more broadly.

Uploaded brains is just one example of a futuristic vision of AI, and science fiction books are filled with plenty more. Nevertheless, it is difficult to imagine forms of AI to which one could not raise important objections. New kinds of techniques can overcome objections to legitimacy raised in the past, but will often meet new concerns. Each time a barrier to legitimacy is crossed by new techniques, there is a possibility that the *z*-axis will expand in that the use of those techniques will be legitimate in circumstances where use of earlier techniques was not. I doubt that we can conceive of an autonomous system that will overcome *all* objections to legitimacy, but I may be proven wrong. Based on the objections raised to date, it would seem that this would require machines to do more than merely *mimic* or *predict* human judges. Rather, people may need to be convinced that machine judges share a similar consciousness to themselves and are oriented to their moral and historical understandings.

VI. Conclusion

Disentangling current practice, technical capacity and legitimacy, with a focus on differences between and within subfields of AI, helps ground debates about AI. Along the *x*-axis, one might observe successes and failures in the implementation of new legal technologies and critique the balance between applications that make large-scale litigation more efficient versus applications that enhance access to justice. Along the *y*-axis, one can look at new techniques introduced over time, each bringing different affordances and limitations that need to be understood when considering the bounds what they make possible. More recently, there has been a growth in scholarship around where the limits of legitimacy may lie, particularly when delegating judicial powers of the state to automated systems. We thus have an evolving picture of the *z*-axis, but the solid is relatively flat.

Straining visualisation further, one might use opacity of the solid generated to signal the *extent* of automation (actual, possible and legitimate) in a given domain. Hybrid systems with machine and human components can then be imagined within the same solid – more opaque where human components are minimal (human-on-the-loop) and more transparent where humans retain greater

[86] K Brennan-Marquez and S Henderson, 'Artificial Intelligence and Role-Reversible Judgment' (2019) 109 *Journal of Criminal Law and Criminology* 137
[87] *Buck v Bell*, 274 US 200 (1927).

control over the system as a whole. Here, legitimacy can grow, for example through automated "pre-instance" decisions overseen by independent human appeal mechanisms offering a full hearing and human judgment.[88]

Barriers that seem insurmountable along all three axes may (or may not) be overcome in the future as technology continues to evolve. However, it is unlikely that future history will play out as a straight path towards the legal singularity. More likely, new technologies will lead to both progress and error, sometimes expanding what is available, what is possible and what is appropriate and legitimate in ways that are both evolutionary and revolutionary. But there are many thresholds to cross, and it is hard to imagine a system that would render law fully computable without changing the nature of law itself.

[88] Wu, above n 18.