

Regulation *by* AI?

The judge's robes are a deep black, though subtle touches of colour complement the national emblem dominating the courtroom wall. Red symbolizes revolution; golden stars rising over the Tiananmen Gate signify the unity of the people under the Party's leadership. Until the turn of the century, judicial officers wore military uniforms – the Supreme People's Court sits at the apex of the legal system but below the Communist Party. By appearance, this judge would not have even been in law school back then. Appearances can be deceiving, of course, since her generic face and simple hairstyle were designed by computer scientists. The avatar's lips move as the synthesized voice asks in Mandarin, 'Does the defendant have any objection to the nature of the judicial blockchain evidence submitted by the plaintiff?'

'No objection,' the human defendant responds.

The video of the pre-trial meeting at Hangzhou's Internet Court, released in late 2019, is part propaganda, part evangelism. Courts were identified as one of the areas ripe for improvement in China's New Generation Artificial Intelligence Development Plan. In a section on social governance [社会治理], it called for the creation of 'smart courts' [智慧法庭].¹ This builds on moves to digitize and standardize litigation across the country, with experiments like those in Hangzhou paving the way for further advances. The avatar can handle online trade disputes, copyright cases, and e-commerce product liability claims.² Hangzhou was chosen because it is the home of Alibaba, enabling integration with

¹ 国务院关于印发新一代人工智能发展规划的通知 [State Council Issued Notice of the New Generation Artificial Intelligence Development Plan] (State Council, Guofa [2017] No 35, 20 July 2017).

² 最高人民法院关于互联网法院审理案件若干问题的规定 [The Provisions of the Supreme People's Court on Several Issues Concerning the Trial of Cases by Internet Courts] (Supreme People's Court, Fasi [2018] No 16, 6 September 2018); Chuanman You, 'Law and Policy of Platform Economy in China' (2020) 39 Computer Law & Security Review 1.

trading platforms like Taobao for the purpose of evidence gathering as well as ‘technical support’.³

Online dispute resolution is not new; eBay has long used it to help parties settle tens of millions of disputes annually.⁴ What is interesting in the Chinese context is the extent to which technology is permeating the court hierarchy not just in mediating small claims but all the way up to the Supreme People’s Court itself.

The Judicial Accountability System [司法责任制] began as a campaign to promote consistency in judgments.⁵ Past efforts had relied on reviews by superiors, but this was deemed impractical and undermined the authority of the judge who heard the case.⁶ AI systems now push similar cases up to a judge prior to a decision, flagging an ‘abnormal judgment warning’ if a proposed outcome departs significantly from past data.⁷ This is part of a suite of technologies that have been adopted, influenced by both the supply of technology companies in China and the demands of a complex and developing legal system. The Wujiang District of Suzhou has trialled a ‘one-click’ summary judgment process, automatically generating proposed grounds of decision complete with sentence.⁸ Other courts are following suit.⁹

Singapore’s Chief Justice has said that developments in China are making ‘machine-assisted court adjudication a reality’. At the same time, he noted, the use of AI within the justice system gives rise to

³ Du Guodong and Yu Meng, ‘A Close Look at Hangzhou Internet Court’, *China Justice Observer* (3 November 2019).

⁴ Pablo Cortés, *The Law of Consumer Redress in an Evolving Digital Market: Upgrading from Alternative to Online Dispute Resolution* (Cambridge University Press 2017) 8; Ethan Katsh and Orna Rabinovich-Einy, *Digital Justice: Technology and the Internet of Disputes* (Oxford University Press 2017).

⁵ 最高人民法院关于统一法律适用加强类案检索的指导意见（试行）[Guiding Opinions of the Supreme People’s Court on Unifying the Application of Laws to Strengthen the Retrieval of Similar Cases (for Trial Implementation)] (Supreme People’s Court, Fafa [2020] No 24, 27 July 2020).

⁶ Cf Margaret YK Woo, ‘Court Reform with Chinese Characteristics’ (2017) 27 *Washington International Law Journal* 241; Junfeng Li et al, ‘Artificial Intelligence Governed by Laws and Regulations’ in Donghan Jin (ed), *Reconstructing Our Orders: Artificial Intelligence and Human Society* (Springer 2018) 61 at 67–71.

⁷ Yu Meng and Du Guodong, ‘Why Are Chinese Courts Turning to AI?’, *The Diplomat* (19 January 2019).

⁸ ‘苏州法院刑案简易程序一键生成 [One-click Generation of the Summary Judgment of the Criminal Case in Suzhou Court]’, *法制日报 [Legal Daily]* (19 June 2017).

⁹ 中国法院的互联网司法 [Chinese Courts and Internet Justice] (Supreme People’s Court of the People’s Republic of China, 2019) 63–65; Yadong Cui, *Artificial Intelligence and Judicial Modernization* (Springer 2020).

a 'unique set of ethical concerns, including those relating to credibility, transparency and accountability'.¹⁰ To this one might add considerations of equity, since the drive towards greater automation is dominated by deep-pocketed clients and ever-closer ties to technology companies, with uncertain consequences for the future administration of justice.¹¹

The impact of AI on the practice of law goes well beyond the scope of this book.¹² This chapter considers the narrower question of whether and how AI systems themselves could support regulation of AI. Insofar as gaps are revealed by the rise of fast, autonomous, and opaque systems, do new rules and new institutions need to be supplemented by new actors in the form of AI regulators and judges?

The first section briefly sketches out past efforts to automate the law. Though AI judges are the most provocative example,¹³ many areas of legal practice and regulation have long been seen as ripe for automation. Despite successes in simple and repetitive tasks, these efforts tended to founder because they were premised on a misconception of law as the mere application of clear rules to agreed facts. In practice, the rules are rarely so clear, and disagreement over facts explains a significant portion of legal disputes.

A more promising approach has been to abandon the goal of thinking 'like a lawyer' and approach legal analysis not as the application of rules to facts but as data. The second section discusses this bottom-up approach to legal analytics, which reveals distinct limitations that are not technical so much as social and political. Even though AI systems are getting ever better at forecasting regulatory outcomes, embracing this across the legal system would represent a fundamental shift from making decisions to predicting them.

Even if regulation by AI generally were possible, then, it is not desirable. Can a special case be made, however, for the regulation of AI systems themselves? If the objection to AI regulators and judges is their inability to appreciate the social context within which legal determinations take place, or legitimacy questions about humans having their fate

¹⁰ Sundaresh Menon, 'Opening of the Legal Year' (Supreme Court, Singapore, 7 January 2019).

¹¹ Seth Katsuya Endo, 'Technological Opacity & Procedural Injustice' (2018) 59 Boston College Law Review 821.

¹² See references cited in the introduction to this book at n 7.

¹³ See, eg, Tania Sourdin, 'Judge v Robot? Artificial Intelligence and Judicial Decision-Making' (2018) 41 UNSW Law Journal 1114; Eugene Volokh, 'Chief Justice Robots' (2019) 68 Duke Law Journal 1135.

determined by statistics, one response is that this need not apply to regulation of AI. The third section discusses how systems could be made to be self-policing. As we have seen, for example, one of the virtues of AI is relative transparency in that simulations can be run with slight variations to look for bias. And, unlike humans, a machine is far more likely to admit to its errors.¹⁴

To the extent that they increase the transparency and human control of AI systems, these developments may be useful. But self-regulation by AI ultimately confronts similar limitations to self-regulation by industry. Though helpful in establishing standards and best practices, red lines will need to be drawn and ultimate oversight conducted by politically legitimate and accountable actors. And, if it is impermissible to outsource inherently governmental functions to fast, autonomous, and opaque machines, enforcement of that prohibition cannot itself be left to those same machines.

9.1 Automating the Law

In the literature on AI and the law, an early theme was that legal practice – viewed essentially as the logical application of rules to established facts – was a strong candidate for automation. Though initially confined to theory,¹⁵ in the 1980s researchers developed prototype systems based on manually created representations of rules in machine-readable form.¹⁶ The enthusiasm was characteristic of the time, preceding as it did one of the ‘AI winters’ that has periodically seen inflated expectations crash against reality.¹⁷

Subsequent decades did see transformations in legal research and document management. These increased lawyers’ access to information and their efficiency in using and sharing it, but did not fundamentally alter their role. Even those encouraging the adoption of technology believed that the inability of AI to emulate human qualities limited its scope for taking on the higher functions of lawyers – the role of judges in particular.¹⁸ As we have

¹⁴ See chapter three, section 3.2.2.

¹⁵ See, eg, L Thorne McCarty, ‘Reflections on TAXMAN: An Experiment in Artificial Intelligence and Legal Reasoning’ (1977) 90 *Harvard Law Review* 837.

¹⁶ See, eg, MJ Sergot et al, ‘The British Nationality Act as a Logic Program’ (1986) 29 *Communications of the ACM* 370.

¹⁷ Anja Oskamp and Marc Lauritsen, ‘AI in Law Practice? So Far, Not Much’ (2002) 10 *Artificial Intelligence and Law* 227.

¹⁸ Richard Susskind, ‘Detmold’s Refutation of Positivism and the Computer Judge’ (1986) 49 *Modern Law Review* 125.

seen in other areas, however, emulating human methods may not be the right or the best approach for reaping the benefits of AI. Autonomous vehicles, to pick an obvious case, are not driven by humanoid robots controlling speed and direction with mechanical hands and feet in substitution of their absent ‘drivers’.

The DoNotPay chatbot, launched in 2015, offered an indication of what might be possible. Written by a seventeen-year-old Stanford student, it followed a series of rules to appeal against parking fines. Similar technology now facilitates other simple tasks from the making of wills to reporting suspected discrimination, yielding efficiencies as well as offering greater access to basic legal services for the wider public.¹⁹ It is also leading to a re-evaluation of what the practice of law means, in the sense of a regulated profession. If a practising certificate or membership of a bar is required to offer legal advice, at what point does an automated system cross that line? Rules-based chatbots do not seem problematic, analogous to a textbook with a flowchart indicating how the law may handle various hypothetical situations. But if an AI system takes in new information, analyses it, and recommends a course of action in a manner that goes beyond the expertise of the programmer, does that become legal advice? Should it be regulated in the same manner as a lawyer?²⁰

These are some of the questions raised by legal tech, a growing area of legal practice.²¹ Having a lawyer sign off on advice is the current solution, much as a partner in a firm might approve a memo drafted in significant part by an intern.²² That was the approach accompanying another high-profile example

¹⁹ Paul Gowder, ‘Transformative Legal Technology and the Rule of Law’ (2018) 68 (Supplement 1) University of Toronto Law Journal 82; Frank Pasquale, ‘A Rule of Persons, Not Machines: The Limits of Legal Automation’ (2019) 87 George Washington Law Review 1, 7–17. It is a stretch, however, to call this automation of certain legal processes ‘AI’ in any meaningful sense.

²⁰ In October 2019, for example, the Hanseatic Bar Association Hamburg successfully challenged Smartlaw, a bot operated by Wolters Kluwer, in the district court of Cologne for operating inconsistently with Germany’s Legal Services Act [*Rechtsdienstleistungsgesetz*]. See further Michael Stockdale and Rebecca Mitchell, ‘Legal Advice Privilege and Artificial Legal Intelligence: Can Robots Give Privileged Legal Advice?’ (2019) 23 International Journal of Evidence & Proof 422; Polly Botsford, Future of Law: Courts Debate Legality of Legal ‘Bots’ (International Bar Association, 11 March 2020).

²¹ Sanda Erdelez and Sheila O’Hare, ‘Legal Informatics: Application of Information Technology in Law’ (1997) 32 Annual Review of Information Science and Technology 367; Jens Frankenreiter and Michael A Livermore, ‘Computational Methods in Legal Analysis’ (2020) 16 Annual Review of Law and Social Science 39.

²² See, eg, Model Rules of Professional Conduct (American Bar Association, 2020), rule 5.3 (responsibilities regarding non-lawyer assistance – though the language of the rule clearly

of technology making inroads into the legal profession, when white-shoe law firm Baker & Hostetler announced that IBM's Ross was joining its bankruptcy practice.²³ Though routinely referred to as a 'robot lawyer', Ross was neither: it was a subscription service lacking any physical form (certainly not a humanoid one), and it did not provide legal advice as such. It was, however, adept at sifting through vast numbers of documents for relevant information in support of the firm's cases.²⁴

Many lawyers long assumed that litigation would be the last part of legal practice to be automated, though the example of China from the introduction to this chapter points to inroads being made there, also. Online dispute settlement has a long history and, for smaller claims in particular, has been embraced not only by online traders like eBay and PayPal but also in the legal systems of Canada and Britain.²⁵

And yet the tsunami of change long forecast by Richard Susskind and others has not (yet) occurred.²⁶

In part this is due to institutional resistance. Lawyers have defended their domain against encroachment by accounting firms and other actors; some view computers as just the next horde to be repelled.²⁷ As a profession, lawyers are also notoriously conservative. Though transactional lawyering must accommodate the needs of business, courtroom procedures retain elements both byzantine and archaic. The Covid-19

assumes that such assistance comes from a 'person'. Cf Ed Walters, 'The Model Rules of Autonomous Conduct: Ethical Responsibilities of Lawyers and Artificial Intelligence' (2019) 35 Georgia State University Law Review 1073; Anthony E Davis, 'The Future of Law Firms (and Lawyers) in the Age of Artificial Intelligence' (2020) 27(1) The Professional Lawyer 3.

²³ Michal Addady, 'Meet Ross, the World's First Robot Lawyer', *Forbes* (12 May 2016).

²⁴ See, eg, Dena Dervanović, 'I, Inhuman Lawyer: Developing Artificial Intelligence in the Legal Profession' in Marcelo Corrales, Mark Fenwick, and Nikolaus Forró (eds), *Robotics, AI and the Future of Law* (Springer 2018) 209 at 226–27; Sergio Alberto Gramitto Ricci, 'Artificial Agents in Corporate Boardrooms' (2020) 105 Cornell Law Review 869, 876. Ross Intelligence announced in December 2020 that it was shutting down operations – defeated not by the limitations of its programming or the open-textured nature of law but by a lawsuit from competitors. Rhys Dipshan, 'ROSS Shuts Down Operations, Citing Financial Burden from Thomson Reuters Lawsuit', *Law.com* (11 December 2020).

²⁵ Richard Susskind, *Online Courts and the Future of Justice* (Oxford University Press 2019).

²⁶ See, eg, Richard Susskind, *The Future of Law: Facing the Challenges of Information Technology* (Oxford University Press 1996); Richard Susskind, *The End of Lawyers? Rethinking the Nature of Legal Services* (Oxford University Press 2008).

²⁷ Chay Brooks, Cristian Gherhes, and Tim Vorley, 'Artificial Intelligence in the Legal Sector: Pressures and Challenges of Transformation' (2020) 13 Cambridge Journal of Regions, Economy, and Society 135, 148.

pandemic of 2020 forced a reassessment of information technology in law firms and the courtroom.²⁸ Much as classes at schools and universities utilized video-conferencing services like Zoom, however, this was a change of medium rather than a transformation of the way in which law is practised.

A second reason why the legal profession has resisted radical change, and may continue to do so, is less self-serving. For it turns out that neither of the assumptions underpinning the hopes for widespread automation – that law is a contained logical system and that facts can be unambiguously established – withstands scrutiny.

9.1.1 *The Inner Illogic of the Law*

A preliminary problem is that legal rules are typically expressed in natural language that may be difficult for a computer to parse. This is a familiar issue in linguistics: humans often interpret language consistently but not logically. Imagine an instruction to go shopping, for example, with the following request: ‘Please buy me a newspaper; and if the store has bananas, buy six.’ A naïve and literal interpretation could lead an autonomous agent to return with six copies of the newspaper. Similarly, the difference between saying that ‘I hunted the bear with my wife’ and ‘I hunted the bear with my knife’ is immediately clear to a human but requires additional information outside the text to make sense.²⁹ Sometimes language may be inherently ambiguous. The statement that ‘I saw the girl with the telescope’ might mean either that the speaker looked through a telescope or that the girl was carrying one.

Advances in natural language processing have overcome many of these difficulties, though statutes and case law may be more challenging than the average text.³⁰ Indeed, the profession of law depends on the ability to

²⁸ Julie Marie Baldwin, John M Eassey, and Erika J Brooke, ‘Court Operations During the COVID-19 Pandemic’ (2020) 45 *American Journal of Criminal Justice* 743. Cf Daphne Yong, ‘The Courtroom Performance’ (1985) 10(3) *The Cambridge Journal of Anthropology* 74.

²⁹ Ian McEwan, *Machines Like Me* (Vintage 2019) 178.

³⁰ See, eg, Livio Robaldo et al, ‘Introduction for Artificial Intelligence and Law: Special Issue “Natural Language Processing for Legal Texts”’ (2019) 27 *Artificial Intelligence and Law* 113; Loïc Vial, Benjamin Lecouteux, and Didier Schwab, ‘Sense Vocabulary Compression through the Semantic Knowledge of WordNet for Neural Word Sense Disambiguation’ (2019) arXiv 1905.05677v3; Boon Peng Yap, Andrew Koh, and Eng Siong Chng, ‘Adapting BERT for Word Sense Disambiguation with Gloss Selection Objective and Example Sentences’ (2020) arXiv 2009.11795v2; Zakaria Kaddari et al, ‘Natural Language Processing: Challenges and Future

charge clients for advice as to how to structure their activities to comply with the law, and advocating on their behalf to enforce it in support of their interests. There may be multiple plausible constructions of a given text – even a carefully drafted one. And until statutes and judgments are written in a manner that can be represented using formal logic, the authoritative text is the original one.³¹

This points to a more fundamental problem, which is that many laws are not reducible to logical representation.³² To be sure, some may be. Road traffic laws, for example, state that exceeding a given speed limit constitutes an offence. Many jurisdictions use speed cameras that automatically record infringements and issue fines. Yet it is telling that these laws – among the most commonly experienced, for much of the population – rarely feature in law school curricula, precisely because they are so clear.³³

Others are not. The tort of negligence, considered in chapter four, is not representable as duty of care plus breach plus causation minus defences equals liability. It explicitly incorporates judgments based on human experience – the famous ‘man on the Clapham omnibus’³⁴ – and notions of reasonableness. In other areas of law, terms such as ‘good faith’ or ‘unconscionability’ are notoriously difficult to define in terms that would be useful to a machine.³⁵ Pretending otherwise is to delegate the interpretive task from the judge not to the machine but to the programmer who establishes its parameters.³⁶ More formally, it is sometimes argued that efforts to treat the law as a logical system susceptible to automation will fail due to the necessary incompleteness of that system – and all such systems.³⁷

Directions’ in Tawfik Masrour, Ibtissam El Hassani, and Anass Cherrafi (eds), *Artificial Intelligence and Industrial Applications* (Springer 2021) 236.

³¹ L Karl Branting, ‘Artificial Intelligence and the Law from a Research Perspective’ (2018) 14(3) *Scitech Lawyer* 32.

³² Cf H Patrick Glenn and Lionel D Smith (eds), *Law and the New Logics* (Cambridge University Press 2017).

³³ Note that many jurisdictions allow ‘reasonable excuse’ as a defence, so perhaps even this example is not so simple.

³⁴ *McQuire v Western Morning News* [1903] 2 KB 100, 109 (Collins MR).

³⁵ See, eg, Mindy Chen-Wishart and Victoria Dixon, ‘Humble Good Faith: 3 x 4’ in Paul Miller and John Oberdiek (eds), *Oxford Studies in Private Law Theory* (Oxford University Press 2020) 187.

³⁶ Francesco Contini, ‘Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings’ (2020) 2(1) *Law, Technology, and Humans* 4, 7.

³⁷ Cf Huws and JC Finnis, ‘On Computable Numbers with an Application to the Alan Turing problem’ (2017) 25 *Artificial Intelligence and Law* 181, 183.

In any case, few legal theorists today would adhere to a strictly formalist position that law can or should be interpreted mechanically. Ronald Dworkin, for example, did hold that there is one correct answer to legal questions – even the difficult ones – but he explicitly rejected the notion that this implied that the answer was reachable by a computer designed by an ‘electronic magician’.³⁸ On the contrary, the difficulty in applying the law is that it is always an exercise in political morality, interpreting the law in its best light on behalf of a community in search of a justification for state coercion.³⁹ Joseph Raz rejected Dworkin’s view of uniquely correct solutions, arguing that judges in such cases are analogous to subordinate legislators, with legal duties to enact particular rules.⁴⁰ The positivist tradition is often seen as the most sympathetic to automation of legal processes, but even HLA Hart held that judges must make choices where existing law fails to dictate that any decision is the ‘correct’ one.⁴¹ Legal realists and critical legal studies scholars, who emphasize the role of judges and the influence of power on the social order, would regard the question of automating the law as so ridiculous to not be worth taking seriously.⁴²

9.1.2 *In Fact*

In his confirmation hearings before the US Senate, Chief Justice John Roberts deflected criticisms of partisanship by quipping that his job was merely ‘to call balls and strikes’. The answer was disingenuous regarding the politicized nature of the court, but Roberts also underestimated the moves to automation in major league sport. In baseball in particular, there have been many calls for umpires to be assisted by a computerized strike zone or replaced entirely. If the role of judges was as simple as determining whether a leather-encased ball passed within a three square

³⁸ Ronald Dworkin, *Law’s Empire* (Harvard University Press 1986) 412.

³⁹ Brian Sheppard, ‘Warming Up to Inscrutability: How Technology Could Challenge Our Concept of Law’ (2018) 68(Supplement 1) *University of Toronto Law Journal* 36, 60.

⁴⁰ Joseph Raz, *Ethics in the Public Domain: Essays in the Morality of Law and Politics* (Clarendon Press 1995) 249–50.

⁴¹ HLA Hart, *The Concept of Law* (3rd edn, Clarendon Press 2012) 273. Cf Abdul Paliwala, ‘Rediscovering Artificial Intelligence and Law: An Inadequate Jurisprudence?’ (2016) 30 *International Review of Law, Computers & Technology* 107.

⁴² Cf Sangchul Park and Haksoo Ko, ‘Machine Learning and Law and Economics: A Preliminary Overview’ (202) 11(2) *Asian Journal of Law and Economics* 15 (adopting a law and economics analysis and concluding that such systems might be treated as expert witnesses but not as substituting for the human judge).

foot zone or not, then they probably should be replaced by machines – it would be both more efficient and more consistent.⁴³

Even if a law appears on its face to be expressed clearly, however – ‘no vehicles in the park’, to pick a well-known example first offered by Hart – how it is to be applied in practice may be less so. We might agree that it covers automobiles, but what about bicycles, roller skates, toy cars?⁴⁴ How about a stroller? Or the statue of a Second World War tank?⁴⁵

The underlying problem is that the strength and the weakness of language is that it is open textured, an idea traceable back to Wittgenstein.⁴⁶ Even when there may be near-universal agreement on many applications of the law, marginal cases will arise. The open-textured nature of language and law has an important connection to time, since future cases may arise that were unknowable by the drafter of a rule. Twentieth-century legislators, for example, could be forgiven for failing to contemplate whether the vehicles prohibited from entering the park include drones.⁴⁷

The need for flexibility in applying the law to particular facts is not merely hypothetical. In the late nineteenth century, the New York State Court of Appeals heard a case in which the plain language of a will and the relevant legislation made clear that the grandson of Francis B Palmer should inherit his estate. Yet the fact that the younger Mr Palmer had poisoned his late grandfather gave them pause. Dworkin uses this example to argue that nearly universal principles of justice may require a departure even from clear textual rules. (The murderer did not get his inheritance.)⁴⁸

Perhaps the strongest illustration of the difficulty of applying law to facts is the market for legal services, in particular litigation. If laws were clearly drafted and easily applied, few disputes would go to court because

⁴³ Jennifer Walker Elrod, ‘Trial by Siri: AI Comes to the Courtroom’ (2020) 57 *Houston Law Review* 1085.

⁴⁴ HLA Hart, ‘Positivism and the Separation of Law and Morals’ (1958) 71 *Harvard Law Review* 593, 607.

⁴⁵ Pierre Schlag, ‘No Vehicles in the Park’ (1999) 23 *Seattle University Law Review* 381; Frederick Schauer, ‘A Critical Guide to Vehicles in the Park’ (2008) 83 *New York University Law Review* 1109.

⁴⁶ Hart (n 41) 124; Ralf Poscher, ‘Ambiguity and Vagueness in Legal Interpretation’ in Lawrence M Solan and Peter M Tiersma (eds), *The Oxford Handbook of Language and Law* (Oxford University Press 2012) 128.

⁴⁷ Michael A Livermore, ‘Rules by Rules’ in Ryan Whalen (ed), *Computational Legal Studies: The Promise and Challenge of Data-Driven Research* (Edward Elgar 2020) 238 at 246–47.

⁴⁸ Ronald Dworkin, *Taking Rights Seriously* (Harvard University Press 1977) 23, citing *Riggs v Palmer*, 115 NY 506 (1889).

rational, well-informed actors would reach the correct conclusion on their own. There would be no need for appellate courts. The reason that cases end up in court is only rarely because one side is objectively and obviously ‘wrong’. This is borne out in practice. Assuming that potential litigants in civil suits make rational estimates of the likely outcome at trial, for example, the individual maximizing decisions of parties should mean that their success rate approaches 50 per cent, regardless of the substantive area of law.⁴⁹ That figure is a limit case only – approached as the standard of decision is clearer, parties’ estimate of the quality of their own cases is more accurate, and the stakes on either side are of similar value. But it finds empirical support.⁵⁰

9.2 Law as Data

Inherent in many of the debates over AI and legal regulation are fundamental differences in the understanding not of AI but of law. If law is understood in a narrowly formalistic way – the blind application of rules to uncontested facts – then processing it through algorithms makes sense, in the same way that it would be inefficient to have regulators or judges doing long division by hand instead of using a calculator.⁵¹ But, to state the obvious, law is not long division. The simplest of cases aside, regulation of behaviour and the resolution of disputes is an inherently agonistic enterprise that involves values and meaning that are necessarily contested.⁵² As Oliver Wendell Holmes famously said, ‘The life of the law has not been logic: it has been experience.’⁵³

Ah yes, the computer scientist might respond. But experience is precisely what machine learning can replicate now.

Indeed, more recent innovations reflect a shift in the approach to the law analogous to the move in AI research towards machine learning. Rather than trying to encode legal rules in fixed systems that can then be applied to sanitized facts – top down, as it were – key achievements have

⁴⁹ George L Priest and Benjamin Klein, ‘The Selection of Disputes for Litigation’ (1984) 13 *Journal of Legal Studies* 1.

⁵⁰ Simon Chesterman, ‘Do Better Lawyers Win More Often? Measures of Advocate Quality and Their Impact in Singapore’s Supreme Court’ (2020) 15 *Asian Journal of Comparative Law* 250.

⁵¹ Mireille Hildebrandt, ‘Law as Information in the Era of Data-Driven Agency’ (2016) 79 *Modern Law Review* 1. For an example of using AI to rethink the notion of legal logic, see Douglas Walton, *Argumentation Methods for Artificial Intelligence in Law* (Springer 2005).

⁵² Jeremy Waldron, ‘The Rule of Law and the Importance of Procedure’ in James E Fleming (ed), *Nomos L: Getting to the Rule of Law* (New York University Press 2011) 3 at 22.

⁵³ Oliver Wendell Holmes, Jr, *The Common Law* (Little, Brown 1881) 1.

been made in analysing large amounts of data from the bottom up. This approach does not seek to answer an individual case, but it does offer a prediction as to the outcome based on past experience.⁵⁴ It represents, as Mireille Hildebrandt observes, a shift ‘from reason to statistics and from argumentation to simulation’.⁵⁵

The turn to AI in this context has proven useful in identifying relevance for the purposes of legal research, contract review, and discovery.⁵⁶ But if extended to regulation and adjudication, it would fundamentally change the task from making a decision to predicting it.⁵⁷ Rather than being part of an ongoing social process in the development of the law, such determinations are more akin to forecasting the weather.⁵⁸ Analytics may provide more information to disputing parties and encourage efficient resolution of disputes while reducing bias and error,⁵⁹ but they could not be a replacement of the judicial function itself.⁶⁰

Indeed, in some jurisdictions the approach has been met with outright hostility. Recall France’s 2019 law prohibiting the publication of data analytics that reveal or predict how particular judges will decide on cases, with a maximum punishment of five years in prison.⁶¹ Though France will likely remain an outlier, AI systems will not replace lawyers or judges in the near term. A more probable scenario is increasing use of AI systems as part of legal services, a partnership sometimes compared to the pairing of humans and machines to play advanced chess, also known as centaur or cyborg chess.⁶²

⁵⁴ Maxi Scherer, ‘Artificial Intelligence and Legal Decision-Making: The Wide Open?’ (2019) 36 *Journal of International Arbitration* 539, 569–71. See, eg, Nikolaos Aletras et al, ‘Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective’ (2016) 2:e93 *PeerJ Computer Science*.

⁵⁵ Mireille Hildebrandt, ‘Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics’ (2018) 68(Supplement 1) *University of Toronto Law Journal* 12, 29.

⁵⁶ See Robert Dale, ‘Law and Word Order: NLP in Legal Tech’ (2019) 25 *Natural Language Engineering* 211.

⁵⁷ Cf Oliver Wendell Holmes, Jr, ‘The Path of the Law’ (1897) 10 *Harvard Law Review* 457, 461 (‘The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by law’).

⁵⁸ Frank Pasquale and Glyn Cashwell, ‘Prediction, Persuasion, and the Jurisprudence of Behaviourism’ (2018) 68(Supplement 1) *University of Toronto Law Journal* 63, 64–65.

⁵⁹ Daniel L Chen, ‘Judicial Analytics and the Great Transformation of American Law’ (2019) 27 *Artificial Intelligence and Law* 15.

⁶⁰ See also the discussion of medical as opposed to legal research in chapter three, introduction.

⁶¹ See chapter three, section 3.3.2.

⁶² See, eg, Rebecca Crootof, ‘“Cyborg Justice” and the Risk of Technological–Legal Lock-In’ (2019) 119 *Columbia Law Review Forum* 233, 243; John Morison and Adam Harkens,

In this context, it is common to draw a distinction between technology assisting in the retrieval of information and that helping in the exercise of judgment.⁶³ The former is analogous to use of a calculator and deemed unproblematic; the latter raises troubling questions about who is exercising discretion. But when the ‘information’ being retrieved goes to the heart of a decision, that distinction may be artificial. As we saw in chapter three, reliance on opaque systems to make recommendations on matters like sentencing is an abdication of the judicial function not because those recommendations may be incorrect but because they are illegitimate. More generally, automation bias raises concerns that human agency may diminish in favour of reliance on the machine.⁶⁴ Even for sophisticated decision-makers, it can be difficult to tell where an algorithm’s ‘nudge’ ends and the accountable individual’s choice begins.⁶⁵

For present purposes, it is sufficient to conclude that AI will continue to transform the legal profession and the role of lawyers – but not to replace them completely. The limits are not so much technical as inherent in the nature of law and the legitimacy accorded to it through political structures in most well-ordered societies.

9.3 Law as Code

Is there, however, a special case to be made for AI playing a larger role in regulating AI itself?

The speed, autonomy, and opacity of AI systems do occasionally give rise to practical and conceptual difficulties for human regulators. In some cases, the response has been to slow them down, as in the case of high-frequency trading.⁶⁶ In others, it has been to ensure the possibility of accountability through requiring that actions be attributable to traditional legal persons – typically the owner, operator, or manufacturer.⁶⁷

‘Re-engineering Justice? Robot Judges, Computerised Courts and (Semi) Automated Legal Decision-Making’ (2019) 39 *Legal Studies* 618, 634–35.

⁶³ Zihuan Xu et al, ‘Case Facts Analysis Method Based on Deep Learning’ in Weiwei Ni et al (eds), *Web Information Systems and Applications* (Springer 2020) 92.

⁶⁴ See chapter three, section 3.1.

⁶⁵ Mariano-Florentino Cuéllar, ‘Cyberdelegation and the Administrative State’ in Nicholas R Parrillo (ed), *Administrative Law from the Inside Out: Essays on Themes in the Work of Jerry L Mashaw* (Cambridge University Press 2017) 134 at 159.

⁶⁶ See chapter one, section 1.2.

⁶⁷ See chapter four, section 4.1.1.

In still others, it has been to call for prohibiting certain activities entirely – most prominently the use of lethal force.⁶⁸

AI does offer means of supporting regulation of AI, although, as discussed in chapter five, the traditional justifications for regulation do not translate easily onto AI systems themselves. In particular, in the absence of AI with legal personality, the targets of regulation are not the AI systems themselves but those who own, operate, and make those systems. That said, the unique features of AI suggest two avenues for a form of self-regulation. First, regulatory objectives can be built into the software itself. Analogous to requirements that privacy values be incorporated into software harvesting personal data, this may be termed regulation by design. Secondly, AI systems allow for interrogation of mistakes and adverse outcomes in a manner not possible with traditional legal actors. This should enable greater transparency concerning errors, but the consequences should also be different than for traditional legal persons. It will be described here as regulation by debugging.

9.3.1 *Regulation by Design*

The idea of incorporating law-compliant behaviour into an AI system may seem self-evident. Autonomous vehicles should comply with traffic laws; algorithms allocating social benefits or recommending loans should not discriminate on the basis of gender or race. But it is possible to go far beyond this.

The notion that regulation can be achieved through design is not new. Though legal scholars often focus on ‘command and control’ approaches, design standards can gather information, set standards, and shape behaviour for regulatory ends.⁶⁹ The usual tools of regulation – commands, incentives, influence – presume the need to compel or persuade human actors (or their corporate proxies) to do or refrain from doing certain actions.⁷⁰ Programmable devices and systems, which include most applications of AI considered in this book, offer the possibility of incorporating regulatory standards directly into their code.

⁶⁸ See chapter two, section 2.2.

⁶⁹ Karen Yeung, ‘“Hypernudge”: Big Data as a Mode of Regulation by Design’ (2017) 20 *Information, Communication & Society* 118, 120. See generally Lawrence Lessig, *Code: Version 2.0* (Basic Books 2006); Mireille Hildebrandt, ‘Saved by Design? The Case of Legal Protection by Design’ (2017) 11 *Nanoethics* 307; Nynke Tromp and Paul Hekkert, *Designing for Society: Products and Services for a Better World* (Bloomsbury Visual Arts 2019).

⁷⁰ See chapter seven, section 7.3.

There are limits. As chapter seven argued, proposals analogous to Asimov's laws of robotics misconceive the nature of law and will never be a complete solution to the regulatory challenges posed by AI systems. But as a restriction on what such systems can do, they point to a promising path forward. As stressed in chapter eight, effective standard-setting will, in some cases, require global rules. Implementing those rules should not rely upon state enforcement alone – to the extent possible, they should be encoded into AI systems themselves. As for the content of those rules, most will be the same that would apply to any product or service. Rather than requiring robots not to murder humans, for example, the prohibition would be against producers making devices that could do so.

Of more interest is how regulation by design might support the two areas identified in chapter seven as potential gaps: human control and transparency.

On human control, building in capability restrictions and a 'kill switch' may sound like obvious design solutions. For the time being, that is true – chapter eight proposed a global agency to support a ban on the creation of uncontrollable or uncontainable AI. As discussed in chapter five, however, in the event that the emergence of a superintelligence becomes more realistic, such constraints could bring about the evil that they are intended to prevent; it may be more prudent to seek to instil alignment with human values instead.⁷¹

In terms of transparency, different degrees are appropriate depending on the type of decision or activity in question. Generally, however, AI systems should be designed to identify themselves as such and in a manner that enables identification of a legal person who is the owner, operator, or manufacturer.⁷² In addition, systems should at a minimum maintain a basic audit trail of how decisions are made.⁷³ This points to the second way in which AI could assist in its own regulation, which is through enabling interrogation of its failures.

9.3.2 *Regulation by Debugging*

When one human kills another, it may give rise to criminal prosecution and lawsuits – these raise legal questions to be resolved. When a machine kills someone, there may be investigation of its owner, operator, or

⁷¹ See chapter five, section 5.3.

⁷² See chapter eight, section 8.2.2(b).

⁷³ See chapter six, section 6.2.2(b).

manufacturer. But with regard to the machine itself, the problem is more likely to be seen as an engineering one. Much as airplane crashes are studied using information from flight data recorders, audit trails in AI systems offer the chance to review how and why errors occurred. If these disclose culpability on the part of the owner, operator, or manufacturer, legal remedies may follow. As for the AI system itself, however, punishment for an error would make no more sense than punishing a plane for its engine failure.⁷⁴

If a system is deemed unsafe, it may be removed from the market; a more likely scenario is that it would be improved. Much as software is now continuously updated with patches as bugs and vulnerabilities are discovered, AI systems operating in the world should be expected to evolve in response to their environment. Market pressure will encourage such updates, but they could also be the subject of regulations or a court order.⁷⁵

Debugging in this way satisfies the aims of regulation at far less cost. Assuming that the improvements do not introduce other errors, it may also be more reliable than traditional regulatory tools if an AI system cannot be tempted once more into deviance. It presumes, of course, a degree of transparency that is unavailable in traditional regulatory settings. If one asks a human driver whether she ran a red light, or a human manager if he discriminated on the basis of race, the answer may be unreliable. Proper audit logs should avoid this problem with respect to AI systems.

This ability to get straight answers also points to another potential strength of such systems, which is that they could be tasked with monitoring themselves. As described in chapter six, two broad theories of oversight are known as ‘police patrol’ and ‘fire alarm’, depending on whether it is conducted through periodic surveys or waiting for problems to be escalated.⁷⁶ AI systems offer a third possibility of self-investigation. This would be more than a regime of self-regulation, as it would not rely on the good faith of actors with incentives to defect.⁷⁷ Provided that the

⁷⁴ See chapter five, section 5.1.2(b).

⁷⁵ Mark A Lemley and Bryan Casey, ‘Remedies for Robots’ (2019) 86 University of Chicago Law Review 1311, 1386–89.

⁷⁶ See chapter six, section 6.1.2.

⁷⁷ See chapter eight, section 8.1. But see Casey Chu, Andrey Zhmoginov, and Mark Sandler, ‘CycleGAN, a Master of Steganography’ (2017) 1712.02950v2 arXiv; Joel Lehman et al, ‘The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities’ (2018) arXiv

instructions were clear, a system could report on its compliance with rules and policies, among other things examining its conduct for bias with a degree of candour not possible with humans.⁷⁸ Problems disclosed in this way would also point to a need to rethink the remedies available – not as sins to be punished but as errors to be corrected.

9.4 The Prospects for Regulation

After the avatar's brief interaction with the parties concludes, the video celebrating Hangzhou's Internet Court shows an interview with its very human vice president, Ni Defeng. 'What we are doing now,' he enthuses, 'you can't understand it as merely improving efficiency. It also speaks to the issue of legal justice. The faster speed – is kind of justice on its own, because justice delayed is justice denied.'

The desire for efficiency and consistency is driving China's push to digitize its court system, with strong endorsement by government as well as the judiciary, and strong support from industry. Though judges themselves remain, for the most part, human, Shanghai's courts are replacing law clerks with AI systems to perform basic legal research – another step in the push to modernize the judicial system through the use of technology.⁷⁹ These developments have been matched by the embrace of computational legal studies in Chinese legal academia. The past decade has seen a more widespread turn to empirical legal studies than in the United States; computational methods are now routinely used in articles published in the top generalist Chinese law journals.⁸⁰

A partial explanation of the greater traction of computational approaches in theory as well as in practice is that China's embrace of the rule of law is more instrumental than its Western counterparts'.⁸¹

1803.03453v1; Tom Simonite, 'When Bots Teach Themselves to Cheat', *Wired* (8 August 2018) (describing AI systems that learned to 'cheat').

⁷⁸ See chapter three, section 3.2.2.

⁷⁹ Sarah Dai, 'Shanghai Judicial Courts Start to Replace Clerks with AI Assistants', *South China Morning Post* (1 April 2020).

⁸⁰ Yingmao Tang and John Zhuang Liu, 'Computational Legal Studies in China: Progress, Challenges, and Future' in Ryan Whalen (ed), *Computational Legal Studies: The Promise and Challenge of Data-Driven Research* (Edward Elgar 2020) 124.

⁸¹ Randall Peerenboom, *China's Long March Toward Rule of Law* (Cambridge University Press 2002) 280–330; Cong-rui Qiao, 'Jurisprudent Shift in China: A Functional Interpretation' (2017) 8(1) *Asian Journal of Law and Economics*; Simon Chesterman, 'Can International Law Survive a Rising China?' (2020) 31 *European Journal of International Law* 1507.

Chinese judges refer to interpretation and the exercise of discretion in the context of ‘judicial measurement’ [裁判尺度], a term without a precise equivalent in the Western tradition but routinely invoked in China with a view to unifying judicial standards.⁸² Judgments at the district and intermediate level tend to be short – a couple of paragraphs stating the facts, an outline of the applicable law and responses to the parties’ arguments, and a decision.

Nevertheless, Chinese judges also express wariness about ‘black box’ decision-making.⁸³ In part this is due to concerns about the accuracy of the outcomes. Initial efforts to train computers on murder cases had to be shelved, for example, because there was an insufficient number of cases and the facts in each varied so greatly.⁸⁴ But it also goes to the trust that underpins the legal system and the rule of law itself.

It remains to be seen whether China represents the future of regulation by AI or its limit case. This chapter has argued that some of the qualities of AI systems that make them hard to regulate through traditional processes may also offer tools to regulate them through new ones. Regulation by design and regulation by debugging suggest ways in which AI systems can be built to comply with the law and tasked with investigating their own biases and failings in a way that most humans would find uncomfortable or impossible.

Yet there are limits to this role. Even if AI systems are more efficient and more consistent than human regulators and judges, that would not justify the handover of their powers more generally.

For the authority of law depends on its processes not only in a formal sense but in a substantive sense also. Regulation, legal decisions, are not mere Turing Tests in which we speculate whether the public can guess if the regulator or judge is a person or a robot. **Legitimacy lies in the process itself, the ability to tie the exercise of discretion to a being capable of**

⁸² 统一裁判尺度 规范法律适用 [Uniform Judgment Standards and Standardize the Application of Law] (Supreme People’s Court of the People’s Republic of China, 12 January 2018). Cf Jiang Na, ‘Old Wine in New Bottles? New Strategies for Judicial Accountability in China’ (2018) 52 *International Journal of Law, Crime and Justice* 74.

⁸³ 郭富民 [Guo Fumin], ‘人工智能无法取代法官的审慎艺术 [Artificial Intelligence Cannot Replace the Prudential Art of Judges]’, 中国法院网 [*China Court Network*] (5 July 2017); Jie-jing Yao and Peng Hui, ‘Research on the Application of Artificial Intelligence in Judicial Trial: Experience from China’ (2020) 1487 *Journal of Physics: Conference Series* 012013, 4.

⁸⁴ Jinting Deng, ‘Should the Common Law System Welcome Artificial Intelligence: A Case Study of China’s Same-Type Case Reference System’ (2019) 3 *Georgetown Law Technology Review* 223, 275.

weighing uncertain values and standing behind that exercise of discretion. Accepting otherwise would be to accept that legal reasoning is not a mix of doctrinal, normative, and interdisciplinary scholarship. Rather, it would come to be seen as a kind of history – the emphasis on appropriate categorization of past practice rather than participation in a forward-looking social project.⁸⁵

As Robert H Jackson, another US Supreme Court judge, once observed: ‘We are not final because we are infallible, but we are infallible only because we are final.’⁸⁶ Many decisions might therefore properly be handed over to the machines. But the final exercise of discretion, public control over the legal processes that regulate our interactions with the world around us, should be transferred only when we are also prepared to transfer political control – when we give up the ballot box for the Xbox.

⁸⁵ Cf Michael A Livermore (ed), *Law as Data: Computation, Text, and the Future of Legal Analysis* (Santa Fe Institute Press 2019).

⁸⁶ *Brown v Allen*, 344 US 443, 540 (1953) (Jackson, J concurring).