THE DEVIL IN THE DETAIL: MITIGATING THE CONSTITUTIONAL & RULE OF LAW RISKS ASSOCIATED WITH THE USE OF ARTIFICIAL INTELLIGENCE IN THE LEGAL DOMAIN

DR. CATRINA DENVIR*

DR. TRISTAN FLETCHER**

MR. JONATHAN HAY***

PROFESSOR PASCOE PLEASENCE****

ABSTRACT

Over the last decade increased emphasis has been placed on the role that artificial intelligence (AI) will play in disrupting the practice of law. Although considerable attention has been given to the practical task of designing a computer to "think like a lawyer," a number of related issues merit further inquiry. Of these, the risks that AI presents to the constitutionally protected procedural and substantive dimensions of justice deserve particular attention. In this Article, we consider the public and private application of AI in the administration of justice and the provision of legal services. We observe that the imposition of AI in certain legal contexts and settings has the potential to silence discourse between actors and agents, subvert the rule of law, and directly and indirectly threaten constitutional rights. In substantiating these observations, we begin in Part I by contextualizing recent developments in legal technology. Tracing the evolution of rule-based AI approaches through to modern data-driven techniques, in Part II we explore how AI systems have sought to represent law, drawing on the domains of: (a) judicial interpretation and reasoning; (b) bargaining and transacting; and (c) enforcement and compliance, and we illustrate how these representations have been constrained by the AI approach used. In Part III we assess the use of AI in legal services, focusing specifically on implications that are posed in respect of the protection of constitutional rights and adherence to the rule of law. Finally, in Part IV we examine the pragmatic challenges that arise in balancing the risks and rewards of AI technologies in the legal domain, and we

* Monash Business School, Monash University, Melbourne, Australia. With thanks to Professor Philip Leith (Queen's University Belfast, UK), Professor Russell Smyth (Monash University, Australia), Professor Nigel Balmer (University College London, UK), Sam Bell, and Seth Nabarro for reviewing and providing helpful comments on earlier drafts. Funding Declaration: This research was supported with a grant made by the Legal Reform Institute.

** Centre for Artificial Intelligence, University College London, London, UK.

*** Judge Business School, University of Cambridge, Cambridge, UK.

**** Centre for Empirical Legal Studies, Faculty of Law, University College London, London, UK.

consider the issues that should shape and that are likely to shape its use. We conclude by proposing the development of a "rule of legal AI" designed to solidify the shared values that ought to govern future development in the field.

I. INTRODUCTION

A. Context

For over five decades researchers have attempted to apply techniques from the field of Artificial Intelligence (AI) to computationally model aspects of legal decision-making.[1] These endeavors have led to the creation of a variety of different systems within and outside of academia, relatively few of which have managed to make the transition from the lab to the marketplace.[2] However, over the last few years, tools employing AI and designed to support legal task completion have come to occupy an increasingly important role in public and private legal services delivery. When combined with substantial growth in the number of legal tech start-ups,[3] recent trends in technology adoption exhibit a new direction for a profession who as recently as 2016 were accused of working practices largely unchanged since the time of Charles Dickens.[4]

For those familiar with the propensity of technology to succumb to reoccurring "hype cycles,"[5] recent developments may be easily dismissed as a passing fad. For others, the influence of AI is not something that should be so easily disregarded; not at least without a more involved examination of the potential consequences that arise when such technologies are given free rein to "weave themselves into the fabric of everyday life until they are indistinguishable from it."[6] Computer systems (particularly those charged with assisting decision-making) present as neutral and value-free, capable of enhancing rather than detracting from the structural integrity of the legal system by minimizing the risks of human error/discretion.[7] However, such systems cannot be judged purely on design, without regard to the seen and unforeseen consequences that arise in implementation.

5. *See* Vicky Harris, *Artificial Intelligence and the Law - Innovation in a Laggard Market?*, 3 J.L. & INFO. SCI. 287, 287 (1992); Edwina L. Rissland et al., *AI and Law: A Fruitful Synergy*, 150 ARTIFICIAL INTELLIGENCE 1, 6 (2003) (discussing Mehl's *Automation in the Legal World* conference paper proposing the use of logic for information retrieval and inference in 1958—only two years after the concept of AI was first defined by McCarthy).

6. Richard E. Susskind, *Expert Systems in Law: Out of the Research Laboratory and into the Marketplace*, *in* PROCEEDINGS OF THE 1ST INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 4 (1987) [hereinafter Susskind, *Out of the Research Labratory*]; Philip Leith, *The Rise and Fall of the Legal Expert System*, 30 INT'L REV. L., COMPUTERS, & TECH. 94, 99 (2016); Anja Oskamp & Marc Lauritsen, *AI in Law Practice? So Far, Not Much*, 10 ARTIFICIAL INTELLIGENCE & L. 227, 227 (2002).

7. *See, e.g.*, Stanford Codex Center for Legal Informatics, STANFORD CODEX, http://techindex.law.stanford.edu (last visited July 1, 2017) (listing 713 entries in June 2017, up from 557 entries recorded in Oct. 2016).

8. Michael Skapinker, *Technology: Breaking the Law*, FIN. TIMES (Apr. 11, 2016), https://www.ft.com/content/c3a9347e-fdb4-11e5-b5f5-070dca6d0a0d) [https://perma.cc/2CQQ-L72Y] (quoting Richard Susskind & Daniel Susskind).

9. Gartner Inc., *Gartner's 2016 Hype Cycle for Emerging Technologies Identifies Three Key Trends That Organizations Must Track to Gain Competitive Advantage*, GARTNER.COM (Aug. 16, 2017), http://www.gartner.com/newsroom/id/3412017 (last visited Mar. 18, 2020).

10. Mark Weiser, *The Computer for the 21st Century*, 265 SCI. AM. 94, 94 (1991).

11. *See, e.g.*, Linda J. Skitka et al., *Accountability and Automation Bias*, 52 INT'L J. HUMAN-COMPUTER STUD. 701, 702 (2000) (discussing the potential for automation to reduce error).

Although AI systems have the potential to guard against human failings (whether those failings relate to capability, efficiency, or self-control), such tools fundamentally change how people approach their work, how systems are organized and operated, and where accountability vests for mistakes.[8] Where decisions involve matters of personal freedom, liberty, or the exercise of rights and responsibilities—as is the case in relation to civil and criminal law—the stakes are vastly higher. Power—whether vested in man or machine—must be accompanied by a commensurate degree of accountability. The values of the justice system that humans strive to uphold are the same values against which AI technology must also be measured. Although there has been considerable attention given to the technical issues of designing a computer to "think like a lawyer," there remains a need for greater reflection as to the theoretical, jurisprudential, and philosophical consequences that accompany this achievement.

In this Article we address this gap, offering a framework by which to understand the application of AI to law and a means by which to assess the impact of this application. We commence in Part I with a series of definitions that contextualize the analysis in the sections that follow. Tracing the evolution of rule-based AI approaches through to modern data-driven techniques, in Part II we explore how AI systems have sought to represent law, drawing on the domains of: (a) judicial interpretation and reasoning, (b) bargaining and transacting, and (c) enforcement and compliance, and we illustrate how these representations have been constrained by the AI approach used. In Part III we assess the use of AI in legal services, focusing specifically on implications that are posed in respect of the protection of constitutional rights and adherence to the rule of law. Finally, in Part IV we examine the pragmatic challenges that arise in balancing the risks and rewards of AI technologies in the legal domain, and we consider the issues that should shape, and are likely to shape, its use. We conclude by proposing the development of a "rule of legal AI" designed to solidify the shared values that ought to govern future development in the field.

We argue that law is a mechanism for balancing competing interests, providing the infrastructure within which those competing interests exist and can be expressed and reconciled. We observe that the imposition of AI in certain legal contexts and settings has the potential to silence discourse between actors and agents, subvert the rule of law, and directly and indirectly threaten constitutional rights. We see this threat as the product of a dualist assumption that it is possible to represent the "law" using technology without that representation being influenced or constrained by its enabling apparatus (technological or otherwise). We propose that understanding and making explicit

12. *Id.* at 701–02.

the model of law being imposed by a particular form of AI is a critical step. One that better positions us to deploy AI in a way that strengthens constitutional values, promotes adherence to the rule of law, and more fairly distributes the associated risks and benefits of new technologies.

B. Definitions

For the purpose of the analysis that follows in Parts II, III, and IV, it is necessary for us to clarify our terminology at the outset. In this section we define our use of the terms "artificial intelligence," "law," and the "rule of law," and briefly outline the assumptions that inform this use.

1. Artificial Intelligence

AI is a broad term incorporating activities involving the design and development of machines, which mimic some of the cognitive functions of the human mind. Whilst AI commonly conjures up thoughts of sentient machines, early experiments in the second half of the 20th century demonstrated the difficulty of replicating the (largely unknown) operations of the human brain to produce "general intelligence."[9] Work within the field has since focused on the development of systems that can perform tasks in relation to specific sub-domains of intelligence, such as learning, problem solving, reasoning, gathering and understanding knowledge, perception, and communication. General AI requires more than the ability to perform a certain action or series of actions. It requires a system to be capable of undertaking decision-making or inference tasks in pursuit of multiple different goals, by drawing on data received via a system of perception such as a camera, sensor, and data packages transferred via networks, keyboard, mice, or microphone.[10] Importantly, AI is considered a moving target with a degree of ephemerality, or as Professor David Miller describes, "whatever machines haven't learned to do yet."[11]

Symbolic approaches to AI development (exemplified by expert systems, logic systems, and information retrieval systems) characterized much of the early work in the field.[12] These approaches focused

---

13. John R. Searle, *Minds, Brains, and Programs*, 3 BEHAV. & BRAIN SCI. 417, 417–18 (1980).

14. STUART RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 25–27 (3d ed. 2010).

15. Jim Dator, *Artilectual Salutations*, 6 J. FUTURE STUD. 87, 89 (2001) (citing Professor David Miller, a robotics specialist at the International Space University and the University of Oklahoma).

16. *See, e.g.*, Michael Aikenhead, *The Uses and Abuses of Neural Networks in Law*, 12 SANTA CLARA COMPUTER & HIGH TECH. L.J. 31, 32 (1996); Andrew Terrett, *Neural*

on the representation of knowledge, with knowledge as to how to reach a goal embedded within the design of the program. More recent efforts have been directed towards sub-symbolic (or "data-driven") approaches such as machine learning, which create a representation of knowledge or latent rules through modeling and analyzing data using statistical methods. In contrast to symbolic approaches, rules are not imposed at the outset deterministically, but are instead discovered (inferred) by mapping outcomes (the correct answer) as a function of inputs (the "facts"). This permits the reverse engineering of rules (which can later be transposed into rule based systems)[13] that capture how the world works in practice and also enables generalization—the ability for systems to extrapolate from the knowledge they have ingested to make predictions from inputs (combinations of data) never before encountered.

For over three decades AI projects have attempted to understand and model legal reasoning. Much of the 1980s was spent speculating as to the possibility that rule-based systems would replace the work of lawyers[14] and some went so far as to argue that this would lead to superior results because "finding chains of consequences in laws, and finding where laws contradict each other, are ideal tasks for computers and are often done poorly by humans."[15] The idea that AI systems can be designed to find chains of consequences in law is in principle quite simple. However, in practice it requires an agreed interpretation of *what the law actually is*—this is far from settled and debate as to how it should be settled exposes longstanding tensions between divergent strands of legal theory.

2. *Law*

Many theories have been offered in pursuit of explaining what the law is—effectively attempting to make sense of what to some would

---

Networks—Towards Predictive Law Machines, 3 INT'L J.L. & INFO. TECH. 94, 110 (1995); David R. Warner, Jr., *A Neural Network-Based Law Machine: Initial Steps*, 18 RUTGERS COMPUTER & TECH. L.J. 51, 53–54 (1992); Jürgen Hollatz, *Analogy Making in Legal Reasoning with Neural Networks and Fuzzy Logic*, 7 ARTIFICIAL INTELLIGENCE & L. 289, 290 (1999); Dan Hunter, *Commercialising Legal Neural Networks*, J. INFO., L. & TECH. at § 1 (1996); John Zeleznikow & Andrew Stranieri, *The Split-Up System: Integrating Neural Networks and Rule-Based Reasoning in the Legal Domain*, *in* PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 185, 185 (1995).

17. *See, e.g.*, Nicholas Frosst & Geoffrey Hinton, *Distilling a Neural Network Into a Soft Decision Tree*, *in* CEX WORKSHOP AT AI*IA 2017 CONFERENCE at § 7 (2017); Qinglong Wang et al., *A Comparative Study of Rule Extraction for Recurrent Neural Networks*, CORNELL U. ARXIV (2018), https://arxiv.org/pdf/1801.05420.pdf [https://perma.cc/B6ZM-FNVM].

18. *E.g.*, Graham Greenleaf, *Legal Expert Systems—Robot Lawyers? An Introduction to Knowledge-Based Applications to Law*, *in* AUSTRALIAN LEGAL CONVENTION at § 1–5 (1989); Leith, *The Rise and Fall of the Legal Expert System*, *supra* note 2, at 94.

19. DONALD MICHIE & RORY JOHNSTON, THE CREATIVE COMPUTER: MACHINE INTELLIGENCE AND HUMAN KNOWLEDGE 54 (1st ed. 1984).

appear as an incoherent patchwork of rights and duties. Efforts to explain "law" often represent it as: the product of a socially accepted process, the proclamations of a sovereign entity, judicial decision-making,[16] the normative expectations held by society as to appropriate behavior, [17] and/or, commands backed by threat. [18] Within this contested space, distilling the complexity of law into a framework for understanding is not an easy task.

Unifying differing explanations inevitably lead us down the path of legal pluralism, in which it is accepted that one or more valid explanations might exist and even interact.[19] On this view, law may be seen as a contestable construct that emerges from the action and reaction between agents and (formal and informal, internal and external, normative and non-normative) structures, existing not just a system of rules intended to control behavior but a system by which to enable access to fair outcomes. For the purposes of this Paper, we take this pluralistic perspective as our starting point. This provides a framework against which we can map the intersection points between AI and the creation, interpretation, application, and enforcement of law as these relate to: (i) judicial decision-making and reasoning, (ii) private bargaining and transacting, and (iii) public enforcement and private compliance.

The legal processes, relationships, rules, and obligations that arise in respect of each of these domains of activity are governed by a theory of order, understood as the "rule of law." This term embodies the

---

20. Roscoe Pound, *Theories of Law*, 22 YALE L.J. 114, 114–16 (1912–1913) [hereinafter Pound, *Theories of Law*]; Brian Leiter, *Legal Realism*, *in* A COMPANION TO PHILOSOPHY OF LAW AND LEGAL THEORY 261 (Dennis M. Patterson ed., 1st ed. 1996) [hereinafter Leiter, *Legal Realism*].

21. EUGEN EHRLICH & WALTER LEWIS, FUNDAMENTAL PRINCIPLES OF THE SOCIOLOGY OF LAW 24 (2001). Ehrlich notes: "it is not an essential element of the concept of law that it be created by the state, nor that it constitute the basis for the decisions of the courts or other tribunals . . ." *Id.* Normative legal interpretation is an expression of Durkheim's 'collective consciousness'—the beliefs and sentiments universal to a people within society. *See also* EMILE DURKHEIM & W. D. HALLS, THE DIVISION OF LABOR IN SOCIETY xix (1997).

22. Austin's command theory perceives law as command backed by threat or sanction—a view in which both the command and the threat are integral criteria for something to be considered "law." Sandra Raponi, *Is Coercion Necessary for Law? The Role of Coercion in International and Domestic Law*, 8 WASH. U. JURIS. REV. 35, 41–43 (2015). Hart offers an alternative position, with the view that enforcement is self-generating: compliance derives from a perception that one is under an obligation to obey rules, leading to social acceptance of rules as constraints upon behavior. *Id.* at 43–46. For a concise overview of Hart and Austin on enforcement, see *id.* at 41–46.

23. Baudouin Dupret, *Legal Pluralism, Plurality of Laws, and Legal Practices: Theories, Critiques, and Praxiological Re-specification*, 1 EUR. J. LEGAL STUD. 296, 297 (2007); John Griffiths, *What is Legal Pluralism?*, 24 J. LEGAL PLURALISM & UNOFFICIAL L. 1, 5 (1986).

constraints that guard against the arbitrary exercise of power and which enable justice to operate in a manner consistent with the principles of liberal democracy.[20]

3. The "Rule of Law"

Exclusive rights, such as the authority to make new formal laws, are bestowed upon only some members of society. These rights amount to a transfer of power that cannot be given without restraint. Governance is therefore subject to the "rule of law," which demands that "the making of particular laws should be guided by open and relatively stable general rules"[21] that promote equality, fairness, predictability, transparency, and accessibility.[22] When taken in conjunction with the absence of corruption and the existence of democratic accountability, the rule of law forms part of a critical triumvirate[23] where it is seen as the "lynchpin for stable government," and "part of the universal duty incumbent on all humanity."[24]

The view of the "rule of law as the law of rules" implies both a constraint on behavior as well as an aspirational framework. As a theory of order rather than of law, adherence to the rule of law does not presuppose a particular outcome, but does govern the process of arriving at that outcome.[25] This is not to say that the scope or purpose of the "rule of law" is agreed—"[l]egal philosophers advance highly nuanced analyses of the rule of law and its contours,"[26] and to this, public agencies often add their own interpretations that focus on

---

24. Randy E. Barnett, *Can Justice and the Rule of Law Be Reconciled*, 11 HARV. J.L. & PUB. POL'Y 597, 622–23 (1988) (discussing the relationship between power and the rule of law); Lord Bingham, *The Rule of Law*, 66 CAMBRIDGE L.J. 67, 84–85 (2007) (discussing the rule of law generally).

25. Joseph Raz, *The Rule of Law and its Virtue*, *in* THE AUTHORITY OF LAW: ESSAYS ON LAW AND MORALITY 210, 213 (1979).

26. Bingham, *supra* note 20, at 5. It should be noted that various authors have sought to describe the "rule of law" in slightly different ways. *See, e.g.*, THE COLLECTED WORKS OF F.A. HAYEK, VOL. XVII: THE CONSTITUTION OF LIBERTY: THE DEFINITIVE EDITION (Ronald Hamowy ed., 2011); Raz, *supra* note 21, at 213; Noel Cox, *Editorial: The Rule of Law as the Product of the Interplay between Potentially Conflicting Conceptions*, 101 ROUND TABLE 299–302 (2012); Michael L Principe, *Albert Venn Dicey and the Principles of the Rule of Law: Is Justice Blind? A Comparative Analysis of the United States and Great Britain*, 22 LOY. L.A. INT'L & COMP. L. REV. 357, 371 (2000); ALBERT VENN DICEY, INTRODUCTION TO THE STUDY OF THE LAW OF THE CONSTITUTION 202–03 (1982).

27. Amir N. Licht et al., *Culture Rules: The Foundations of the Rule of Law and Other Norms of Governance*, 35 J. COMP. ECO. 659, 663 (2007).

28. Cox, *supra* note 22, at 299.

29. Noel B. Reynolds, *Legal Theory and the Rule of Law*, NEWSL. PHIL. & L. (Am. Philos. Ass'n Newark, DE), Spring 2002, at 119.

30. Licht et al., *supra* note 23, at 663–64.

dimensions of policy or economic development.[27] Notwithstanding these differing views (a detailed analysis of which is beyond the scope of this Paper), the rule of law is generally said to prioritize a number of key features, included among which are: stable and prospective rules, an independent judiciary, limits on discretion, observance of the principles of natural justice, accessible courts, and the constraint of power/authority.[28]

In the United States, the principles enshrined in the rule of law find formal expression via the written constitution (and accompanying amendments) which operates to limit government, impose checks and balances, ensure the separation of powers, and guarantee citizens basic rights, such as equal protection and due process. The principles and values at the heart of the rule of law and in respect of which the constitution is silent, also find expression informally outside of a constitutional framework, via statute, common law, and norms of behavior. In jurisdictions where a codified constitution does not exist, such as the United Kingdom, this model of informal expression has given rise to what is known as an "unwritten constitution."[29] For the purposes of the forthcoming analysis, we use "rule of law" to refer to both of these formal and informal expressions so as to accommodate the varying degrees of formality by which they are recognized in various jurisdictions.

II. THE DEVELOPMENT OF LAW MACHINES

Amplified by claims that vast parts of the legal profession can be turned over to machines, the potential of AI and machine learning (ML) in law has been met with both enthusiasm and existential anxiety.[30] Yet for these claims to be plausible, it must be possible for a computer to replicate dimensions of "human legal intelligence." In this section we examine the development of "law machines" in relation to the creation, interpretation, application, and enforcement of law, examining three key areas where progress has been focused: (1) judicial

31. *See, e.g.*, *What is the Rule of Law?,* WORLD JUSTICE PROJECT, https://worldjusticeproject.org/about-us/overview/what-rule-law (last visited Jan. 19, 2018) (explaining that the UN, advances a human-rights centric definition as compared to the World Justice Project).

32. *See, e.g.*, Greenleaf, *supra* note 22, at § 1–5; Leith, *The Rise and Fall of the Legal Expert System*, *supra* note 2, at 94.

33. John Baker, *The Unwritten Constitution of the United Kingdom*, 15 ECCLESIASTICAL L.J. 4, 19 (2013).

34. *See, e.g.*, Chris Weller, *Law Firms of the Future Will be Filled with Robot Lawyers*, BUSINESS INSIDER (July 7, 2016, 12:09 PM), https://www.businessinsider.com/law-firms-are-starting-to-use-robot-lawyers-2016-7); James O'Toole, *Here Come the Robot Lawyers*, CNN BUSINESS (Mar. 28, 2014, 7:16 AM), https://money.cnn.com/2014/03/28/technology/innovation/robot-lawyers/ [https://perma.cc/7U4A-Z3K7].

interpretation and reasoning, (2) private bargaining and transacting, and (3) public enforcement and private regulatory compliance.

A. Judicial Interpretation & Reasoning

Much of the work of lawyers in the field of litigation is directed at predicting what a court might decide when interpreting the law, or when applying the law to a particular factual scenario.[31] Whilst it is not possible to know with absolute certainty how it is that judges reason, a number of theories have been offered.[32] These describe legal reasoning as the analysis of legal rules enshrined in legal sources using methods that are: logical and deductive; analogical and comparative; or subjective and discretionary, and these approaches to explaining reasoning have underpinned the development of a range of different AI systems intended to replicate the process of legal interpretation.

1. Rules & Rule-Based Systems

For formalists, the role of the judge is not to make law but rather to bring clarity to the existing law (whatever its source) by engaging in textual analysis to understand the plain meaning of the words intended to bring that law into effect. Thus, legal reasoning (and the laws elucidated through this process) is borne out of logical deduction in which conclusions necessarily follow from the premises. For example:

If A, then B

A ∴ B

The internal logic of the proposition is valid: it holds true no matter what the values of A and B.

If the law is interpreted in the way that the formalists propose, then the task of the judge is merely to apply the appropriate rules of interpretation. The law is largely stable, and discretion is limited. Given the same inputs (legislation, existing cases, a contract), same training (and legal education is broadly consistent within jurisdictions), and task (interpreting the meaning of a particular source) it should be possible for a legal professional to anticipate the likely

---

35. Dru Stevenson & Nicholas J. Wagoner, *Bargaining in the Shadow of Big Data*, 67 FLA. L. REV. 1337, 1339 (2015).

36. *E.g.*, E.C. Lashbrooke, Jr., *Legal Reasoning and Artificial Intelligence*, 34 LOY. L. REV. 287, 287 (1988); *see also* NEIL MACCORMICK, LEGAL REASONING AND LEGAL THEORY (2003); STEVEN J. BURTON, AN INTRODUCTION TO LAW AND LEGAL REASONING (3d ed. 2007) (an introductory book meant to teach law students how to mimic the legal analysis used by judges and lawyers); Wilson Huhn, *The Stages of Legal Reasoning: Formalism, Analogy, and Realism*, 48 VILL. L. REV. 305, 307 (2003); Edward H. Levi, *An Introduction to Legal Reasoning*, 15 U. CHI. L. REV. 501, 501 (1948); Brian Leiter, *Positivism, Formalism, Realism*, 99 COLUM. L. REV. 1138, 1138 (1999).

interpretation a judge will reach. When cast in this light, the law is the product of an objective, detached reasoning process, in which fixed rules are applied to source materials.

This being the case, it is not difficult to see how one might develop a rule-based system that simulates deductive reasoning (if A then B, if B then C, ∴ if A then C). The development of logic programming languages in the 1980s made this possible, for unlike procedural languages in which variables and relationships had to be defined in advance, logic programming (implemented using languages such as Prolog) permitted inferences to be drawn by setting a series of logic conditions. These conditions form the attributes a solution must possess, rather than the steps that must be taken to achieve a solution. Through a computational process called unification, the system identifies all possible solutions capable of satisfying underlying logic. An example of this process is detailed in Figure 1.

In procedural programming ‘Who’ could not equal ‘bob’, unless Who = ‘bob’ had been previously specified within the system. In logic programming it is possible to pass in uninstantiated arguments such as ‘Who’. Via a process called ‘unification’, the interpreter will attempt to deduce what ‘Who’ should equal using previously defined rules/clauses.[33] This process is shown below.

Prolog Clause	**Meaning**
1. Lawyer (john)	*John is a Lawyer*
2. Lawyer (james)	*James is a Lawyer*
3. Colleagues (john, mary)	*John and Mary are colleagues*
4. Colleagues (james, mary)	*James and Mary are colleagues*
5. Colleagues (susan, x) <- Lawyer (x) & Colleagues (x, mary)	*Susan works with any Lawyer that works with Mary*
Query	**Meaning**
I. Colleagues (susan, james)?	Are Susan and James colleagues?
II. -? Colleagues (james, Who)	Who is a colleague of James?

QUERY: ‘colleagues (susan, james)?’

GOAL 1: Match ‘colleagues’ on the left side of any clauses, preceding the parenthesis.
MATCHED: Clause 3, 4 and 5

GOAL 2: ‘susan’ and ‘james’ in the correct ordered position within the parenthesis for clauses 3, 4 and 5.
MATCHED Clause 5 – ‘susan’ is in the correct position & ‘james’ matches ‘x’ (which denotes ‘anyone’).

GOAL 3: Match the right side of Clause 5 – ‘Lawyer (james) & Colleagues (james, mary)’
MATCHED Lawyer (james) is proven by Clause 2 and Colleagues (james, mary) is proven by Clause 4.

OUTPUT: Colleagues (susan, james)? = TRUE

Similarly, the query ‘-? Colleagues (james, Who)’ would match ‘Colleagues’ first, and then seek to match james in the first position in the open parenthesis. As clause 4 matches the query ‘Colleagues (James,) the system would return the value Who = mary.

Figure 1. Logic Programming: An Example Using Prolog

The development of rule-based systems involved experts in various domains attempting to distill their reasoning processes into a chain of deductive logic, in which the conclusions follow from the premise (“modus ponens”).[34] This logic was replicated computationally with a knowledge base (a set of facts) and a rules-engine (a set of rules that describes what role the facts have on an outcome, e.g., IF, THEN). These symbolic logic approaches, initially implemented using Prolog, were, according to some, very suitable in law, owing to the fact that “the law is well documented; its provisions are written down, and

33. ALLEN HUSTLER, PROGRAMMING LAW IN LOGIC, RESEARCH REPORT CS-82-13 (Department of Computer Science ed., 1982).
38. *E.g.*, Lashbrooke, *supra* note 23, at 304.

where they are not, decisions in previous cases are recorded for future reference."[35] This led one group of scholars to develop the British Nationality Act System—among the first efforts to translate legal reasoning into an expert system.[36] Developed in 1986, the program formalized 150 rules associated with the obtainment of British citizenship.[37] Table 1 provides a brief overview of the logic applied in the program and the (lay) translation of statute to logic, demonstrating how cascading rules guide a determination of entitlement to citizenship.

Table 1. British Nationality Act Program—Sample Rules & Logic[38]

Statute	**Sub-Rules**	**Logic**
1) A person born in the United Kingdom after commencement shall be a British citizen if at the time of the birth his father or mother is – (a) a British citizen; or (b) settled in the UK.	[A] X acquires British Citizenship by section 1.1 [B] X is born in the UK at T (Time) [C] T is after commencement of the Act [D] Y is parent of X [E] Y is a British citizen at T [F] Y is settled in the UK at T	A is true if [B *and* C *and* D *and* [E *or* F]] are true.

The use of logic programming to emulate aspects of legal reasoning was, at least initially, met with enthusiasm[39] and following the British Nationality Act System, a number of other prototypes were developed in academia and industry in North America, the UK, and Europe.[40] Explaining the enthusiasm surrounding symbolic approaches, Hunter observed that these approaches were fairly easy for lawyers to understand and fitted into the existing normative pedagogical framework, because "law schools teach law as a type of symbolic manipulation, and some go so far as to introduce classes on logic and argument."[41] Others

39. M. J. Sergot et al., *The British Nationality Act as a Logic Program*, 29 COMM. ACM 370, 383 (1986).

40. *Id.*; Edwina Rissland, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, 99 YALE L.J. 1957, 1967–68 (1990) (discussing the evolution of expert systems and situating BNA as one of the first expert systems in law).

41. *Id.*

42. Kowalski, *supra* note 14, § 2.1.

43. With respect to enthusiasm, see Richard Susskind, *Expert Systems in Law: A Jurisprudential Approach To Artificial Intelligence And Legal Reasoning*, 49 MOD. L. REV. 168, 168 (1986) [hereinafter Susskind, *A Jurisprudential Approach To Artificial Intelligence And Legal Reasoning*]; MICHIE & JOHNSTON, *supra* note 15, at 54. For a more critical view of developments at the time and afterwards, see Greenleaf, *supra* note 14, § 1–5; Leith, *The Rise and Fall of the Legal Expert System*, *supra* note 2, at 100.

44. Susskind, *Out Of The Research Laboratory*, *supra* note 2, at 1–8; Richard E. Susskind, *The Latent Damage System: A Jurisprudential Analysis*, *in* PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 23, 23–24 (1989).

45. Dan Hunter, *Looking For Law In All The Wrong Places: Legal Theory and Legal Neural Networks*, *in* JURIX : THE FOUNDATION FOR LEGAL KNOWLEDGE SYSTEMS (1994).

(notably Leith) urged caution in overstating the potential of expert/logic systems in law and underestimating the challenges.[42]

Whilst logic systems such as Prolog enabled the development of AI software, they were considered by some to be unsuitable for knowledge representation.[43] Further, criticism was directed at both the practicality and desirability of what was derisively referred to as "mechanical jurisprudence"[44]—translating legal reasoning to mere symbol manipulation,[45] and the very "notion of a clear rule [of interpretation] which . . . [judges] can apply without further thought."[46] These criticisms arose from recognition that in the real world A and B take on different values, and A ∴ B can easily fail as a result of simplicity.[47] These views sought to counterbalance the claims made by others that progress in the field was held back by an absence of legal knowledge engineers, or a lack of understanding of the nature of law and legal reasoning.[48] Instead, the problems were said to stem from the fact that the law and legal reasoning are contested rather than settled concepts. Systematizing one particular model of interpretation amounts to the imposition of certain assumptions about the law that were not—nor are they now—universally agreed upon.

The real world limitations of simple deduction are easily seen in practice. Take for instance, the following example provided by Article I, Section 5, Clause 4 of the U.S. Constitution, which states, "Neither House, during the Session of Congress, shall, without the [prior] Consent of the other, adjourn for more than three days, nor to any other place than that in which the two Houses shall be sitting."[49] This could be translated into a logic statement as follows:

> House of Representatives adjourns for more than three days without the consent of the Senate
>
> ∴ House of Representatives has acted unconstitutionally

Here the conclusion follows from the premise in terms of the internal logic, but this logic does not illuminate what constitutes "[prior consent]," or "adjournment." Further, what is simple is not necessarily correct; the logic statement does not explain whether

---

46. Philip Leith, *Logic, Formal Models and Legal Reasoning*, 24 JURIMETRICS J. 334, 356 (1984); Philip Leith, *Fundamental Errors in Legal Logic Programming*, 29 COMPUTER J. 545, 545 (1986).

47. Thomas F. Gordon*, Some Problems with Prolog as a Knowledge Representation Language for Legal Expert Systems*, 3 INT'L REV. LAW, COMP. & TECH. 1, 52 (1987).

48. *See* Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605, 623 (1908).

49. Leith, *Fundamental Errors in Legal Logic Programming*, *supra* note 42, at 545.

50. *Id.* at 547.

51. Leith, *Logic, Formal Models and Legal Reasoning*, *supra* note 42, at 339.

52. Susskind, *Out Of The Research Laboratory*, *supra* note 2, at 2.

53. U.S. CONST., art. 1, § 5, cl. 4.

"[prior] consent" is intended to qualify: (a) adjournments longer than three days *or* adjournments to another place, or (b) adjournments to another place *that* exceed three days in duration.

Expert systems addressed these issues though processes such as backtracking, which allows chains of reasoning enabled rules to be used in a definitional manner.[50] When applied in practice, this results in an increasingly complex set of linked rules, nudging us closer to a definition for each vague term and syntactical ambiguity that exists. However, whilst such an approach may be effective in respect of the vast majority of cases where definitions are settled (what Hart referred to as the "hardcore of standard"), those which arguably are less likely to support differing interpretations (whether pernicious or genuine), what of the "penumbra" in which interpretation is required: how is it that "hard cases" are handled?[51] Where the definitional quality of rules become exhausted, the practical limitations of this form of deductive reasoning, and the difficulty associated with the indeterminacy of language and the (often deliberate) vagueness of legal language are brought into focus.[52]

Admittedly, these challenges have not gone ignored. The limitations of deductive rules of interpretation as described above formed the focus of Gardner's early work conducted at Stanford, looking at what happens when the rules run out, and leading her to develop a computational model capable of distinguishing between hard and easy cases.[53] Developers since have also attempted to accommodate the non-fixed and open-texture nature of law in a range of different ways, including taking cues from expert judgment,[54] deferring to the user,[55] conceptual models, and fuzzy logic.[56] McCarthy's US Taxman

---

54. Leith, *Logic, Formal Models and Legal Reasoning*, *supra* note 42, at 351.

55. H.L.A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593, 607 (1958).

56. Paul Conway, *Syntactic Ambiguity*, L. & JUST. FOUND. NSW, Mar. 14, 2002, at 35–37. In relation to pernicious ambiguity, see Lawrence Solan, *Pernicious Ambiguity in Contracts and Statutes*, 79 CHI.-KENT L. REV. 859, 860–63 (2004).

57. ANNE VON DER LIETH GARDNER, AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING 14–16 (1987).

58. Susskind, *A Jurisprudential Approach To Artificial Intelligence And Legal Reasoning*, *supra* note 39, at 176.

59. T.J.M. Bench-Capon, *Deep Models, Normative Reasoning and Legal Expert Systems*, *in* PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 37, 40 (1989); Greenleaf, *supra* note 14, at 9.

60. Classical logic has it that an object is either part of a class (= 1) or not part of a class (= 0). Fuzzy logic avoids this demarcation, proposing that an object can be part of a class to a certain degree (between 0 and 1). This allows us to represent imprecise concepts and handle uncertain or incomplete information. Ambiguous and overlapping concepts like "short," "medium," and "long" can be represented in fuzzy logic sets, allowing us to take account of the lack of definitiveness of the concepts. *See* Trevor Bench-Capon, *Neural Networks and Open Texture*, *in* PROCEEDINGS OF THE 4TH INTERNATIONAL CONFERENCE

project, for example, "proposed a *prototype-plus-deformation* model for representing ambiguous terms. The prototype [element] stor[ed] the default meaning," whilst the deformations were structured, mapped synonyms.[57] Nevertheless, where definitions are not settled, or ambiguity or vagueness arises, developers find themselves having to make determinations about the law that may not reflect its machinations in the real world. Such systems must also incorporate the ancillary contextual or cultural knowledge that acts to inform the legal rules governing a situation.[58]

These challenges are not unique to law. Encoding the ability of clinicians to perform general reasoning on the basis of previously unseen combinations of patient characteristics or symptoms has also thwarted the progress of logic systems in medicine.[59] To some degree these issues can be avoided in law, though not without bringing new challenges to the fore. For example, imposing a particular interpretation can be avoided by deferring to the judgment of the user, allowing the user to make a call as to whether the conditions for a certain definition have been met. Yet, deferring to the user risks creating a system that conceptually maps legal reasoning, but does not perform it. Conversely, enshrining the expertise (and the judgment) of a lawyer within the system ascribes this expertise a level of definitiveness and authority that conflicts with the way in which expertise is brought to bear in reality. Lawyers advocate for an interpretation that best serves their client, relying on axioms (knowledge of some previous truth), induction, and/or deduction to justify this conclusion. Disputes arise because experts are able to justify competing logical interpretations, and therefore expert systems must do more than just provide a heuristic answer. Instead, they must search the solution space to corral

---

ON ARTIFICIAL INTELLIGENCE AND LAW 292, 292 (1993); *see also* Hollatz, *supra* note 12, at 289–300 (providing useful exploration of fuzzy logic in respect of German jurisprudence); Célia Da Costa Peia et al., *Combining Fuzzy Logic and Formal Argumentation for Legal Interpretation*, 10 *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 49, 55–58 (2017) (providing an implementation example).

61. Seth R. Goldman et al., *Precedent-Based Legal Reasoning and Knowledge Acquisition in Contract Law: A Process Model*, *in* PROCEEDINGS OF THE 1ST INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 210, 220 (1987) (demonstrating that in ascertaining whether a simple contract exists knowledge of social roles, knowledge of idioms and an understanding of the notion of "promise" and of the meaning of "agreement" is required.); L. Thorne McCarty, *Reflections on TAXMAN: An Experiment in Artificial Intelligence and Legal Reasoning*, 90 HARV. L. REV. 837, 888 (1977).

58. Goldman et al., *supra* note 57, at 211.

59. Peter Szolovits, *Artificial Intelligence and Medicine*, *in* ARTIFICIAL INTELLIGENCE IN MEDICINE 16 (Peter Szolovits ed., 1982) (exemplifying this challenge with the following hypothetical: a doctor hears that a patient, who they know works in a feed store, is experiencing upper body pain; the doctor may naturally expect the pain to be caused by heavy-lifting while working, while an AI agent may not have access to this external information, or may be unable to deduce when an occupation may have a medical impact by hard-coded rules).

all possible answers and from this present an argument for each that undermines opposing positions and strengthens the position of the answer in focus.[60] In this regard, law shares some of the strategic elements of games, such as chess; yet unlike chess, in law the range of possible moves and the rules governing these moves are not static or fixed, as the body of case law makes clear.

2. *Precedent & Case-Based Systems*

In the face of competing interpretations, it is argued that judges must have recourse to other forms of reasoning.[61] Within an adversarial system based on precedent, determining whether an A is a B also demands **analogical/inductive** logic in which opposing yet plausible interpretations are reconciled with reference to past cases, as follows:

Case c1 is similar to Case c2

Proposition p is true in case C1

∴ p is true in c2.

In order to automate comparative and analogical reasoning processes, a system must operationalize the concept of "similarity between cases"—determining whether a case fits within a particular class by exploring the factors (key issues) on which a decision of similarity is said to turn. It is not just the existence of similarity of factors, but the polarity (favoring a plaintiff/favoring a defendant) of factors that matter.[62] Similarity speaks to individual factors, as well as to case outcomes as a whole. The following example helps illustrate the point.

In U.S. Trademark Law, "consumer confusion" represents a central consideration for the court in determining whether trademark infringement has occurred.[63] In establishing whether "consumer confusion" exists or is likely to arise, a district court must conduct a multifactor analysis based on the factors set out by that circuit. According to the Seventh Circuit, the multifactor test is intended "as a heuristic device to assist in determining whether confusion exists."[64]

60. KEVIN D. ASHLEY, ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS: NEW TOOLS FOR LAW PRACTICE IN THE DIGITAL AGE 3 (2017).

61. Leith, *Logic, Formal Models and Legal Reasoning*, *supra* note 2, 352–56 (discussing the relevance of inductive and analogical reasoning); Leith, *The Rise and Fall of the Legal Expert System*, *supra* note 2, at 101–03.

62. *Id.* at 76.

63. Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1582 (2006) ("The overriding question in most federal trademark infringement litigation is a simple one: is the defendant's trademark, because of its similarity to the plaintiff's trademark, causing or likely to cause consumer confusion as to the true source of the defendant's goods?").

64. Sullivan v. CBS Corp., 385 F.3d 772, 778 (7th Cir. 2004).

Each circuit has gone on to adumbrate its own formulation of the test, so whilst the Federal Circuit draws on thirteen factors, other circuits, such as the Second, use only eight.[65] In determining whether the defendant's trademark is causing or likely to cause consumer confusion as to the true source of the defendant's goods, the Second Circuit has identified eight factors of importance:[66]

- The Strength of the Plaintiff's Mark (F_1);
- The Similarity of Plaintiff's and Defendant's Marks (F_2);
- The Proximity of Plaintiff's and Defendant's Products (F_3);
- The Likelihood That One of the Parties Will Bridge the Gap (F_4);
- Evidence of Actual Confusion (F_5);
- Defendant's Intent (F_6);
- The Quality of Defendant's Goods (F_7);
- The Sophistication of the Consumers (F_8).

Courts have also advised that the correct approach to analyzing these variables and likelihood of confusion as a whole is to adopt a "global appreciation" of all relevant factors and the contribution they might make to confusion, treating them as interdependent in the sense that a lesser degree of similarity of one factor may be compensated for by a greater degree of similarity of the other(s).[67] This permits analysis in respect of outcomes for individual factors, and analysis in respect of the case outcome as a whole.

Firstly, it is possible to analogize between cases on the basis of the specific sub-factors (e.g., F_1–F_8) to deduce what has led to a conclusion of confusion in previous cases and what that might mean in the context of a new dispute. However, this is more complex than might be assumed. That courts have previously found that the marks "PROZAC" and "HERBROZAC"[68] demonstrate similarity, but that "POLAROID" and "POLARAD"[69] do not demonstrate

---

65. Beebe, *supra* note 63, at 1582–83 (citing Polaroid Corp. v. Polarad Elecs. Corp., 287 F.2d 492, 495 (2d Cir. 1961)).

66. *Polarad*, 287 F.2d at 495.

67. *See, e.g.*, Thane Int'l, Inc. v. Trek Bicycle Corp., 305 F.3d 894, 901 (9th Cir. 2002); Scott Fetzer Co. v. House of Vacuums Inc., 381 F.3d 477, 485 (5th Cir. 2004); Shakespeare Co. v. Silstar Corp. of Am. Inc., 110 F.3d 234, 242 (4th Cir. 1997); Soc'y of Fin. Exam'rs v. Nat'l Ass'n of Certified Fraud Exam'rs Inc., 41 F.3d 223, 228 n.15 (5th Cir. 1995).

68. Eli Lilly & Co. v. Natural Answers, Inc., 233 F.3d 456, 469 (7th Cir. 2000).

69. Polaroid Corp. v. Polarad Electronics Corp., 287 F.2d 492, 495 (2d Cir. 1961).

similarity, suggests that a finding of similarity of marks is not merely a matter of the linguistic "edit distance" between the marks.[70]

It is then also possible to ascertain how the combination of factors (F_1–F_8) contribute to the overall outcome. However, the factors are interdependent and do not contribute an equal weight to an overall finding in favor of the Plaintiff (P) or Defendant (D). That is to say that a strong finding in favor of a plaintiff with respect to one factor may compensate for the fact that many other factors were found in favor of the defendant. Outcomes are not cumulative, nor have lawyers been said to think in this way. As Ashley and Rissland explain: "Experts in domains like the law simply do not reason in terms of weighting schemes. In fact[,] in the legal domain, any reasoner that based an opinion or course of action upon a purely numerical scheme would be highly suspect."[71]

As a result, it is not possible to conclude that a case win is simply the product of more factors being found in favor of the plaintiff then being found in favor of the defendant. Analogizing between cases may instead mean exploring overlaps between cases on the basis of factors, as shown in Figure 2 below.

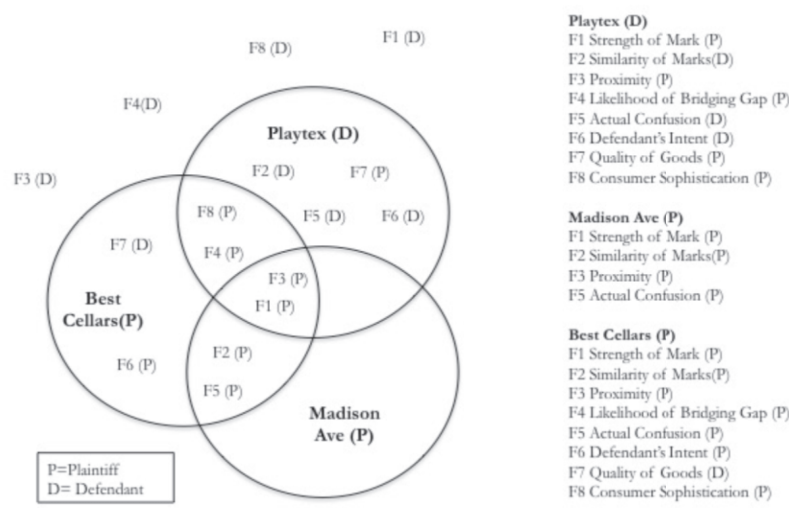

Figure 2. Likelihood of Confusion (F_1–F_8)
Factor Overlap for Three Second Circuit Likelihood of Confusion Cases.

70. A measure of difference computed with reference to the number of operations required to transform one word into another. *See* DANIEL JURAFSKY & JAMES H. MARTIN, SPEECH AND LANGUAGE PROCESSING § 3.11 (2d ed. 2009).

71. Kevin D. Ashley & Edwina L. Rissland, *Waiting on Weighting: A Symbolic Least Commitment Approach*, *in* AAAI PROCEEDINGS 239, 239 (1988).

Figure 2[72] draws together overlaps between three cases decided in the Southern District of New York in accordance with Second Circuit precedent: *Playtex Products v. Georgia-Pacific Inc.*,[73] *Madison Avenue Caviarteria, Inc. v. Caviaria.com*,[74] and *Best Cellars Inc. v. Grape Finds at Dupont, Inc*.[75] It reveals most obviously in regards to *Playtex Products*, that decisions are not made simply by counting the number of factors that favor the Defendant or the Plaintiff.[76] The *Playtex Products* decision shares a number of factor outcomes in common with both *Madison* and *Best Cellars* where a finding of "consumer confusion" was found (P). That factors F_1 and F_3 overlap with both *Madison* and *Best Cellars* might give a lawyer scope to argue that the case has more in common with those where likelihood of "consumer confusion" has been found. Conversely, the fact that *Playtex Products* shares F_8 and F_4 with only *Best Cellars* and both *Madison* and *Best Cellars* share decisions of F_2 and F_5 in favor of the Plaintiff whereas *Playtex Products* does not, gives a lawyer room to argue that the case is sufficiently different from those in which the Plaintiff has won.

Case-based reasoning addresses the notion of legal decision making as analogical, ordinal, and top-down. Outcomes in key cases heard in higher courts, (in principle) bind determinations in lower courts. For this reason, analogizing or distinguishing between the facts of a current case and that of the leading cases replicates the way in which cases are argued in court. This approach has led to the production of a number of systems, including HYPO (on which the methodology conveyed in Figure 2 is based), which compares sets of overlapping dimensions between cases with respect to U.S. Trade Secret Law.[77] It has also led to probabilistic variations on the aforementioned case-based and rule-based approaches, as exemplified by the Shyster program developed in Australia.[78]

Nevertheless, in practice, such systems are of questionable utility. Similarities between cases are likely to be well known to domain

---

72. For the purposes of simplifying this example, the following explanation ignores the order (year) in which the cases were decided, although it is recognized that this will inevitably have an impact upon comparability between cases/factors.

73. Playtex Prods., Inc. v. Ga-Pac. Corp., No. 02 Civ. 7838(HB), 2003 WL 21929706 (S.D.N.Y. Aug. 12, 2003).

74. Madison Ave. Caviarteria, Inc. v. Caviaria.com, No. 04 Civ. 00493 RO, 2004 WL 744481 (S.D.N.Y. Apr. 7, 2004).

75. Best Cellars, Inc. v. Grape Finds at Dupont, Inc., 90 F. Supp. 2d 431 (S.D.N.Y. 2000).

76. *Playtex Prods.*, 2003 WL 21929706, at *3–6.

77. Edwina L. Rissland & Kevin D. Ashley, *A Case-Based System for Trade Secrets Law*, *in* PROCEEDINGS OF THE 1ST INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 60, 61–63 (1987); Stefanie Brüninghaus & Kevin D. Ashley, *Generating Legal Arguments and Predictions from case Texts*, *in* PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 65, 67–69 (2005).

78. *See generally* JAMES POPPLE, A PRAGMATIC LEGAL EXPERT SYSTEM (1996).

experts, and not easily understandable to lay users.[79] Whilst these systems prompt compelling intellectual questions about the nature of legal reasoning, their broader commercial value is unclear. Furthermore, case-based reasoning relies on the presumption of consistency and the view that differences in case outcomes derive from the nuances of reasoning. That decisions *ought* to be based on deductive or analogical reasoning is not to say that they are, even if outwardly this appears to be the case.[80] The easiest way to avoid being committed to deductive rules alleged to prevail in an instance where "if P then Q" would ordinarily apply, is to determine that the facts do not permit a determination of P. This amounts to a Wittgensteinian re-writing of the rules such that "no course of action could be determined by a rule, because every course of action can be made out to accord with the rule."[81] But if reasoning by deduction or analogy does not account for the interpretive process, what does?

3. *Discretion: Modeling Inputs & Outputs*

The formalist view that judges *interpret* rather than *make* law demands that any reasoning must stem from a plain reading of relevant legislation, legal codes, and the case law rules of common law and equity derived from previous judgements. Constraining inputs to a specific range of sources operates to restrict a judge from projecting his or her own opinions, values, or beliefs onto the process of decision-making. Thus, in theory, the positivist/formalist system of law is a closed system in which new legal interpretations are grounded in and arise from existing interpretations. However, if the existing legal material is indeterminate because it can always be interpreted in conflicting and contradictory ways, what drives the decision to adopt one interpretation over another?

On this view, legal interpretation becomes less a form of bounded reasoning (of the deductive/analogical variety) and more a process of subjective political decision-making. This arguably casts doubt on the idea that the law is a closed system, or that judges are as textual as formalists suggest. That judges profess to be following rules or precedent does not guarantee that this is the case. There may be other types of reasoning that influence a decision, even if (by virtue of

---

79. Perhaps explaining why some systems were originally designed to assist law students in learning the basics of legal reasoning, for example, Ashley and Aleven's CATO program. ASHLEY, *supra* note 60, at 3; Vincent Aleven & Kevin D. Ashley, *Evaluating a Learning Environment for Case-Based Argumentation Skills*, *in* PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 170, 174 (1997).

80. *See* Beebe, *supra* note 63, at 1583–85 (discussing this in the context of trademark law).

81. LUDWIG WITTGENSTEIN, PHILOSOPHICAL INVESTIGATIONS § 201 (G.E.M. Anscombe trans., 3d ed. 1986). The relevance of this paradox in this context has also been noted by Leith, *Logic, Formal Models and Legal Reasoning*, *supra* note 42, at 340.

convention in which judges feel compelled to be seen to be following the deductive or analogical "rules") this reasoning does not appear in the *ratio decidendi*.[82] Hence, from a realist perspective, judicial interpretation amounts to the creation of law because it is only via interpretation that legal meaning is formally settled. Judicial discretion, when coupled with the indeterminacy of law, provides scope for reasoning to be influenced by a range of extra-systematic factors, including (but not limited to): (i) the purpose of the law/intent of legislators, (ii) fairness of outcome,[83] and (iii) the broader social/policy/political/economic/legal impact of the decision.[84]

On this view, the inevitable subjectivity of judicial interpretation can only be revealed through the quantitative or qualitative study of case outcomes in which a broad range of potential influences are tested. This testing may involve qualitative techniques (e.g., textual analysis that looks within decisions for evidence of motivations and considerations that go beyond the prevailing logical or functional criteria ordinarily applied), though it is more commonly associated with quantitative techniques, notably the use of descriptive and inferential statistics, and more recently, analysis via ML. These methods provide a way of testing the relationship between case inputs and case outcomes, producing evidence to counter the view that the process of interpretation is "scientific" in nature.

The application of these methods to legal decision-making in order to test the validity of theoretical propositions reveals the influence of a wide range of extraneous factors on judicial decision-making. One such influence is mental fatigue, with research proposing that successive decision-making taxes an individual's executive function and mental resources, increasing the tendency to simplify decisions by accepting the status quo.[85] In the case of repeated judgments or decisions this may explain the increased similarity of decisions[86] and the increased reliance on intuitive decision-making[87] previously observed. Other research has documented more obscure influences,

82. *See generally* Beebe, *supra* note 53; William E. Boyd, *Law in Computers and Computers in Law: Lawyer's View of the State of the Art*, 14 ARIZ. L. REV. 267, 284–85 (1972).

83. MacCormick termed this "Everyday Logic" to mean "that which makes sense" primarily because, in contrast to formal logic, everyday logic aligns with patterns of individual and social belief that give rise to expectations of normative behavior. *See* NEIL MACCORMICK, LEGAL REASONING AND LEGAL THEORY ii (2003).

84. *See supra* note 82.

85. Mark Muraven & Roy F. Baumeister, *Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle?*, 126 PSYCHOL. BULL. 247, 247–48 (2000).

86. *See, e.g.*, Shai Danziger et al., *Extraneous Factors in Judicial Decisions*, 108 PROC. NAT'L ACAD. SCI. U.S.A. 6889, 6890 (2011) (discussing a relationship between the order of a parole decision (before or after a food break) and the favorability of the decision to an applicant).

87. Anastasiya Pocheptsova et al., *Deciding Without Resources: Resource Depletion and Choice in Context*, 46 J. MKT. RES. 344, 353 (2009).

including correlations between the weather, current affairs, and case outcomes in asylum seeker status adjudications.[88] Whilst certain personal characteristics, including the political bias,[89] ethnicity, age, gender, and educational background of judges have been shown to predict judicial decisions in other studies, perhaps explaining why appointments to the U.S. Supreme Court are so contentious.[90]

Modeling law using statistical methods, such as regression analysis, involves reducing a particular outcome (for example: a legal decision, the likelihood of litigation, a settlement amount, or the time taken to progress through court) to a mathematical function. A hypothetical example helps illustrate the point. Assume that we have in our possession a large number of case files on clinical (professional) negligence claims. Insurers are required to record potential settlement or court-awarded amounts as liabilities on corporate accounting records, and lawyers to whom this estimation task may be assigned are judged on accuracy of their assessments. Experienced lawyers commonly review case files and suggest an appropriate financial figure that they believe the court would produce and then calculate from this the amount likely to encourage a claimant to settle rather than to pursue litigation. They draw on professional experience to inform these calculations but these subjective (albeit informed) predictions are not always accurate, nor does accuracy necessarily increase alongside experience.[91]

Although this process of reasoning is not well understood, it is not unthinkable that case features/characteristics seen by lawyers as relevant in informing a settlement/judgement figure could be articulated and encoded in a symbolic, rules-based system. Yet this would explain only one lawyer's perceptions of relevance, and would not account for actual values recorded, values that may be the product of several negotiations rather than the acceptance of a first offer. There is benefit in standardizing the approach taken to predict settlement amounts, by modeling the data in our case files to quantitatively explore the relationship between various case features (variables) and outcome amount. The same would also be true of variations on our hypothetical

---

88. Daniel L. Chen & Jess Eagel, *Can Machine Learning Help Predict the Outcome of Asylum Adjudications?*, *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 237, 238 (2016).

89. *See, e.g.*, CASS R. SUNSTEIN ET AL., ARE JUDGES POLITICAL? AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY 150 (2006) (discussing the empirical evidence supporting politicized judicial decision-making).

90. *See, e.g.*, Justin D. Levinson et al., *Judging Implicit Bias: A National Empirical Study of Judicial Sterotypes*, 69 FLA. L. REV. 63, 68–69 (2017); David S. Abrams et. al., *Do Judges Vary in their Treatment of Race?*, 41 J. LEGAL STUD. 347, 350 (2012); Thomas J. Miles & Cass R. Sustein, *The Real World of Arbitrariness Review*, 75 U. CHIC. L. REV. 761, 765 (2008).

91. Stevenson & Wagoner, *supra* note 22, at 1346–47.

if we were dealing with other forms of legal decision-making, such as winning a case, evaluating parole risk, and/or calculating other damages.

In our insurance hypothetical, descriptive statistics help us identify (among other things) how many cases go to court, how many settle at first offer, and/or the number of cases in which the initial settlement amount is substantially different from the final settlement amount. Inferential statistics involves deriving underlying probability distributions that may help assess differences between groups/test hypotheses or make estimates. This allows us to extrapolate inferences beyond our sample; to examine whether certain case or claimant features (e.g., age of claimant, postcode, type of injury) are associated with higher/lower settlement amounts, and to determine how confident (probabilistically speaking) we should be in concluding that these patterns indicate a real relationship rather than mere coincidence. Assuming our sample is representative, the relationships revealed in our analysis can be generalized beyond the cases we have examined. It is also possible to link the inputs to the outputs, so as to determine the size of effect generated by different model parameters (variables/features). That is, to examine the effect of claimant postcode or income on settlement amount, as compared to the effect of injury type of settlement amount. Doing so helps to reveal the features (variables) that are really driving settlement outcomes.

The scientific model of inquiry on which the statistical model rests, requires an analysis to be directed towards testing a particular hypothesis. In order to predict settlement amounts, we need an idea of the variables/features of a case that we believe to be driving settlement figures, and these go on to form the basis of our hypothesis. However, if the interpretation of law is as inchoate a process as the preceding analysis and the competing explanations of legal reasoning suggest, we may not have a clear idea of the causes producing an effect. This circumstance calls for ML approaches, as these require no hypothesis testing or preconceived theory.[92] Rather than testing select factors, ML models (which typically rely on vastly larger datasets than that used for hypothesis-driven quantitative work, with a far larger number of variables), effectively take the process of inference one step back.[93] Through the ingest of testing and training data, and the mapping of inputs to outputs, only those features (variables) shown to significantly contribute to an outcome (or to the accuracy of predicting the right outcome when comparing testing/training material) are included. These features inform the

92. *See e.g.*, Lyria Bennett-Moses & Janet Chan, *Using Big Data for Legal and Law Enforcement Decisions Testing the New Tools*, 37 UNSW L.J., 643, 648 (2014).

93. *Id.*

development of a post-hoc explanatory theory,[94] though the rational construction of such a theory may prove difficult, for reasons discussed further below.

ML methods (all of which have a statistical genesis and many of which are employed in hypothesis testing approaches as well) are distinguished based on their learning style.[95] Supervised or unsupervised are the most prominent paradigms, though there are many methods that fall outside this dichotomy (e.g., reinforcement learning) or which take elements from both (e.g., semi-supervised methods).[96] Supervised learning methods are intended to produce predictions in respect of unseen data points (e.g., cases where the case features are known but the outcome is not) by mapping existing data.[97] They associate a set of inputs (factors) with another set of outputs (often termed labels), with the implication being that some human effort is required to establish and connect to the inputs.[98] Unsupervised learning is more often focused on finding patterns within sets of data and attempting to attach a meaning to these patterns *a posteriori*, for example, clustering claimants based on their age, profession, income and location, assuming they must share other similarities, and consequently grouping them together for the purposes of litigation strategy.[99]

Both supervised and unsupervised approaches to ML, as well as hybrid (semi-supervised) approaches, have been employed in the academic setting. Chen and Eagel used supervised learning to create a model to predict the outcome of U.S. asylum adjudications using a set of input data.[100] Wongchaisuwat's U.S. Litigation Model identified factors capable of predicting whether a patent is likely to be litigated against and when this might happen, employing methods from both the supervised learning domain (using an ensemble (voting) classification method) in conjunction with a clustering (unsupervised learning)

---

94. *Id.*

95. *See generally* KEVIN P. MURPHY, MACHINE LEARNING: A PROBABILISTIC PERSPECTIVE (2012).

96. *See e.g.*, RICHARD SUTTON & ANDREW G. BARTO, REINFORCEMENT LEARNING: AN INTRODUCTION (2018).

97. CHRISTOPHER M. BISHOP, PATTERN RECOGNITION AND MACHINE LEARNING (2006).

98. These labels can be real valued numbers (e.g., a dollar value when attempting to predict a likely settlement amount) or categories (e.g., whether a claimant is likely to: accept first settlement offer, accept later settlement offer, or refuse to settle). Murphy, *supra* note 95.

99. For an example of some methods for classification and analysis of multivariate observations, see J. MacQueen, *Some Methods for Classification and Analysis of Multivariate Observations*, *in* PROCEEDINGS OF THE FIFTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY 281–297 (1967).

100. Chen & Eagel, *supra* note 88, at 238.

method to improve performance.[101] Whilst Bochereau employed a canonical supervised learning method to ascertain whether the Conseil d'Etat would annul or confirm a bylaw on the basis of input variables relating to regulations, bylaws, factual, and normative standards.[102]

Whilst the range of problems to which ML has been applied in law has varied, common to all implementations irrespective of subject matter is the need for quality data. Input data for ML can take many different forms, drawing from the settlement example above, this might include real valued numerical (scale), categorical, or nominal data. Whilst complex linguistic or semantic constructs can be simplified for the purpose of analysis (e.g., recording the severity of claimant injury as a point on a scale of 1 (minor) to 10 (severe)), there may be some work required to get to data into this format. This is because data must be "structured" in a way that allows for its inclusion in statistical/machine learning models, necessitating organization in the form of columns and rows (or equivalent).[103] This requirement produces certain challenges in law where much of the data used is "unstructured," taking the form of natural language.

For this reason, recent advances in Natural Language Processing (NLP) (the computational processing or "structuring" of language drawing on rule-based, statistical, and ML techniques), which have enabled text-based information to be incorporated into the ML/ statistical modeling process, are of clear significance for fields such as law. NLP uses data-driven and rule-based techniques to translate linguistic meaning into numerical meaning.[104] One such

---

101. Papis Wongchaisuwat et al., *Predicting Litigation Likelihood and Time to Litigation for Patents*, *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 257, 257 (2017).

102. Laurent Bochereau et al., *Extracting Legal Knowledge by Means of a Multilayer Neural Network Application to Municipal Jurisprudence*, *in* PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 288, 290 (1991).

103. For example, SQL data commonly takes the form of relational tables, whereas non-SQL data, which is often scripted as JSON, does not present with the traditional column and row format, though nonetheless, it is still capable of assuming that form.

104. *See generally* Julie Beth Lovins, *Development of a Stemming Algorithm*, MECHANICAL TRANSLATION AND COMPUTATIONAL LINGUISTICS 22–31 (1968). Rule-based methods have fallen out of fashion for achieving complex tasks (e.g., detecting sentiment from text), but are still commonly used during pre-processing. For example, using a stemming algorithm, to 'undo the rules' that allow for all words in a body of text to be reduced to their roots (e.g., 'walking' is mapped to 'walk', 'willingness' is mapped to 'will'), so as to gain topic insight. Mitchell Marcus et al., *Building a large annotated corpus of English: The Penn Treebank*, UNIV. PA. SCHOLARLY COMMONS, Oct. 1993; William Schuler et al., *Broad-Coverage Parsing Using Human-Like Memory Constraints*, 36 COMPUTATIONAL LINGUISTICS 1, 313–330 (2010). Similarly, grammatical rules may be hand-crafted by a linguistic expert, Eric Brill, *A Simple Rule-Based Part of Speech Trigger*, *in* PROCEEDINGS OF THE THIRD CONFERENCE ON APPLIED NATURAL LANGUAGE PROCESSING, 152–55 (1992), with the goal of performing 'part-of-speech' (POS) tagging where each word in a text is assigned a label denoting its grammatical role (noun, verb, adverb et cetera). *See* Table 2 in Laura Chiticariu

implementation of NLP is feature extraction, which identifies and extracts key content from written materials. For example, when using summaries of a patient's discharge reports to deduce a measure of severity of injury, having a feature that reflects the presence of the words "induced coma" might be a good indicator of whether the patient suffered extensive trauma.[105] More complex word embedding methods use (unsupervised) ML to quantify the context of a word based on the surrounding text, with semantically similar words assigned mathematically "close" values-in vector space.[106] This consistency allows for mathematical manipulation (Berlin - Germany + France is approximately equal to Paris[107] and King - Man + Woman is approximately equal to Queen[108]) and can be expanded to incorporate sentences.

These and more complex NLP techniques, such as deep learning, have been used in a range of legal tasks associated with the process

---

et al., *Rule-based Information Extraction is Dead! Long Live Rule-based Information Extraction Systems!*, *in* PROCEEDINGS OF THE 2013 CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING 827, 829 (2013). However, specification of the rules may be time-consuming and requires dependence on expert linguists. It can also give rise to the potential for a high level of inter-annotator disagreement which operates to reduce scalability, whilst irregularities in language are difficult to address and not uncommon. Dat Quoc & Son Bao Pham, *Ripple Down Rules for Question Answering*, SEMANTIC WEB (2015), http://www.semantic-web-journal.net/content/ripple-down-rules-question-answering-0 [https://perma.cc/2SSD-8XHG]. Further, large numbers of rules may be used to capture such irregularities, but interactions between rules may cause complications. CHRISTOPHER D. MANNING, INTRODUCTION TO INFORMATION RETRIEVAL 293–320 (2008). In contrast, ML-based NLP does not require the same level of manual tuning; instead one hopes that the training data is sufficient, and the model employed is apt to capture the relevant linguistic rules. Nonetheless, rule-based feature engineering remains in use for some ML modelling algorithms.

105. Note that while feature engineering often uses simple rules, ("is the phrase 'induced coma' present in the text?"), the output of this rule is not the final output we are interested in (the patient's severity of injury), but forms an input to an ML model, which predicts the final output. This example was modified from a hypertension example provided in Vijay N. Garla & Cynthia Brandt, *Ontology-Guided Feature Engineering for Clinical Text Classification*, 45 J. BIOMED. INFORMMATICS 992, 995–98 (2012).

106. *See* Miguel Kakanakou, *Build and Visualize Word2Vec Model on Amazon Reviews*, BEEXPERT (Sept. 10, 2017), http://migsena.com/build-and-visualize-word2vec-model-on-amazon-reviews/; Christian S. Perone, *Voynich Manuscript: Word Vectors and t-SNE Visualization of Some Patterns*, TERRA ICOGNITA (Jan. 16, 2016), http://blog.christianperone.com/2016/01/voynich-manuscript-word-vectors-and-t-sne-visualization-of-some-patterns/ [https://perma.cc/ABQ9-UWYY]; Kaspar Beelen, *Visualizing Parliamentary Discourse with Word2Vec and Gephi*, ON HISTORY (Aug. 5, 2015), https://blog.history.ac.uk/2015/08/visualizing-parliamentary-discourse-with-word2vec-and-gephi/ [https://perma.cc/C3BB-HXFA].

107. Tomas Mikolov et al., *Distributed Representations of Words and Phrases and their Compositionality*, *in* 2 PROCEEDINGS OF THE 26TH INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS 3111, 3115 (2013).

108. Ekaterina Vylomova et al., Take *and* Took*,* Gaggle *and* Goose*,* Book *and* Read*: Evaluating the Utility of Vector Differences for Lexical Relation Learning*, *in* PROCEEDINGS OF THE 54TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS 1671, 1671 (2016).

of legal interpretation: to predict whether a new bill will become law,[109] to accelerate legal review tasks,[110] to identify deontic modalities in financial regulations,[111] to facilitate an originalist interpretation of legal language,[112] to improve the readability of legislative sentences,[113] to create question and answer dialogues enabling naturalist interrogation of Supreme Court decisions,[114] to automate text summarization,[115] to derive legal knowledge[116] and argument schematics[117] from legal texts, to quantify latent and manifest linguistic similarities between legal documents,[118] and to automate the review of national legislation to identify implementation of European Union (EU) directives.[119]

Using the aforementioned NLP techniques to structure data and ML techniques to expose correlations between different variables and outcomes, it is possible to test the objectivity of judicial decision-

---

109. John J. Nay, *Predicting and Understanding Law-Making with Word Vectors and an Ensemble Model*, 12 PLO ONE, May 10, 2017, at 12.

110. Ngoc Phuoc An Vo et al., *Experimenting Word Embeddings in Assisting Legal Review*, *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 189, 192 (2017).

111. Paul Buitelaar et al., *Classifying Sentential Modality in Legal Language: a use Case in Financial Regulations, Acts and Directives*, 10 *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 159, 159 (2017).

112. *See, e.g.*, Lawrence M. Solan & Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 BYU L. REV. 1311, 1312–13 (2017); Lee J. Strang, *How Big Data Can Increase Originalism's Methodological Rigor: Using Corpus Linguistics to Reveal Original Language Conventions*, 50 U.C. DAVIS L. REV. 1181, 1212–23 (2017); Nathan Kozuskanich, *Originalism, History, and the Second Amendment: What Did Bearing Arms Really Mean to the Founders?*, 10 J. CONST. L. 413, 415–16 (2008).

113. *See, e.g.*, Michael Curtotti et al., *Machine Learning for Readability of Legislative Sentences*, *in* PROCEEDINGS OF THE 15TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 53, 53 (2015); L. Karl Branting et al., *Automated Drafting of Self-Explaining Documents*, *in* PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 72, 72 (1997).

114. Jose Gabriel Lopes et al., *Question/Answer Dialogues for Interfacing a Database with Supreme Court Decisions*, *in* PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 255, 255 (1997).

115. Claire Grover et al., *Automatic Summarisation of Legal Documents*, *in* PROCEEDINGS OF THE 9TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 243, 243 (2003); Ben Hachey & Claire Grover, *Automatic Legal Text Summarisation: Experiments with Summary Structuring*, *in* PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 75, 77 (2005).

116. Vassilis Konstantinou et al., *Can Legal Knowledge be Derived from Legal Texts?*, *in* PROCEEDINGS OF THE 4TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 218, 224 (1993).

121. Katarzyna Budzynska & Serena Villata, *Argument Mining*, 17 IEEE INTELL. INFORMATICS BULL. 1, 1–2 (2016).

118. Erich Schweighofer et al., *Information Filtering: The Computation of Similarities in Large Corpora of Legal Texts*, *in* PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 119, 119 (1995).

119. Rohan Nanda et al., *A Unifying Similarity Measure for Automated Identification of National Implementations of European Union Directives*, *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 149, 150 (2017).

making, and to expose subjective, circumstantial, or ostensibly irrelevant factors influencing legal outcomes. However, whilst this type of quantitative analysis is intended to inform legal reasoning, it is not always credited with this accomplishment. Quantitative empirical methods of understanding decision-making or interpretation (linguistic or otherwise) have attracted criticism for failing to illuminate certain aspects of legal reasoning.[120] To those who approach legal reasoning from a different theoretical position, such findings may be discounted as entirely obvious, and the value they add in systematically validating what to some is "obvious," is often overlooked. Conversely, findings may be viewed as counter-intuitive and nonsensical—not just because they contradict widely held views as to the objective and logical nature of legal reasoning.

Variables are included in a hypothesis-led model because existing theories drawn from law or related fields suggest a relationship between that input and output. As such, where a relationship is shown to exist, the reasoning underpinning the theory serves as the starting point by which to explain the actuality. However, the bigger the available data, the greater risk of severing the theory-data connection,[121] particularly when employing ML techniques in which variable selection is not theory-led. This may see certain variables enhance predictive accuracy even though such variables do not enhance our understanding of judicial decision-making.[122]

Issues translating quantitative findings in a legally meaningful way are magnified where the relationship between inputs and outputs is not just difficult to explain but also difficult to disentangle, as is often the case when using Big Data or NLP and drawing on thousands of variables/features. Claims that the hugely complex black box models produced in ML are of limited utility in respect of understanding phenomena are not unique to legal decision-making: the Defense Advanced Research Policy Agency (DARPA) has created a research

---

120. *See* Ashley & Rissland, *supra* note 60, at 240; Ashley, *supra* note 64, at 100 (discussing the Hypo system, Ashley reports that "legal factor weights are sensitive to the particular context . . . judges and attorneys do not argue about the weight of legal factors in quantitative terms . . . legal domain experts do not agree what the weights are, and combining positive and negative weights numerically obscures the need for arguing about the resolution of competing legal factors").

121. OSONDE OSOBA & WILLIAM WELSER IV, AN INTELLIGENCE IN OUR IMAGE: THE RISKS OF BIAS AND ERRORS IN ARTIFICIAL INTELLIGENCE 17–18 (2017); OSONDE A. OSOBA & PAUL K. DAVIS, *An Artificial Intelligence/Machine Learning Perspective on Social Simulation New Data and New Challenges* 11 (DARPA, Working Paper No. 1213, 2018).

122. Chen and Eagel's paper operates as a case in point, with their predictive model able to correctly classify 82% of cases on the basis of a mix of variables, including case factors, judge factors, the weather, and the news cycle. *See* Chen & Eagel, *supra* note 71, at 238. Whilst incorporating weather in the model enhances its predictive accuracy, weather is not an obvious influence on judicial behavior and serves only to confuse rather than enhance an understanding of judicial decision-making. *Id.*

group called Explainable Artificial Intelligence[123] with interpretability as its key focus.[124] For this reason, models, which clearly articulate how important each input is in decision-making, are favored over those where the model internals are so complicated that this information is not retrievable.[125] However, there are circumstances in which the features relied on by an ML algorithm will be of secondary interest. If, for example, the sole function of an ML algorithm is to classify cases into group A or B, then assuming the ML algorithm replicates human classification with sufficient precision (by drawing on a set of human-annotated A and B cases), addressing the theory-data gap may be unnecessary.[126] Nevertheless, acknowledging and understanding these issues at the outset is critical and this importance is magnified as natural language processing is used to yield increasingly complex models.

As the preceding section makes clear, it is possible (albeit with some caveats) to represent hypothetical and actual judicial interpretation and decision-making as a function of cascading rules, overlapping precedent, and/or the influence of subjective bias represented by statistical patterns in data. The examples discussed above highlight the methods for doing so. Yet whilst most scholarship in the field of AI and law has focused on "black letter law"—that is, "the basic principles of law generally accepted by the courts and/or embodied in the statutes of a particular jurisdiction,"[127] this is only one part of the AI and law picture. The embodiment of law in AI systems extends beyond judicial interpretation, encompassing both the bargains struck between private parties and the public and private enforcement governing those bargains, as the following section reveals.

B. Private Bargaining & Transacting

Within the framework of the law, individuals have autonomy to engage in private ordering of disputes, creating, interpreting, and enforcing their own permutations of "law," and rights/responsibilities through bargaining and agreement. Private bargaining and

---

123. *See generally* David Gunning, *Explainable Artificial Intelligence Program Update November 2017*, DARPA/120 (2017), https://www.darpa.mil/attachments/XAIProgramUpdate.pdf [https://perma.cc/9MSN-LXSR].

124. *Matt Turek*, *Explainable Artificial Intelligence (XAI)*, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY, https://www.darpa.mil/program/explainable-artificial-intelligence [https://perma.cc/X7QC-HJ5A].

125. Gunning, *supra* note 123.

126. Osobra and Davis, *supra* note 117, at iii (The theory-data gap refers to the "mismatch between measurable data streams and meaningful explanatory theories to frame the data.").

127. John Zeleznikow et al., *Bargaining in the Shadow of the Law–Using Utility Functions to Support Legal Negotiation*, in PROCEEDINGS OF THE 11TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 237, 237 (2007).

transacting have given rise to two key inter-related manifestations of AI: (1) tools designed to reduce the transaction costs of bargaining, and (2) tools to simulate the process of bargaining itself.

1. Reducing Bargaining Transaction Costs

The private ordering (or application of law) between parties is motivated by a common goal—pursuing agreement as to a transaction or exchange. This agreement is reached (or not) by bargaining, incentivized by utility[128] gain for both parties, and formalized via some official or unofficial means (e.g., written contract, handshake, verbal agreement, performance). As with legislation and case law, the rights and responsibilities that derive from a private agreement are also subject to interpretation; occasionally by a judge in the context of a dispute, but more often by the parties themselves in the process of bargaining, and for the purpose of informing estimates regarding the utility of an agreement. This interpretation may be influenced by cultural, social, or community expectations in a way that reveals a normative basis to law.[129] It may also be influenced by presumptions about other legal sources, what they are seen to protect/enforce, and what they might conceivably demand of a party in the event of non-performance. These perceptions are learned in a social context[130] which gives rise to the potential for misperception and misalignment between parties. Indeed, the possibility of conflict between parties as to interpretation is not so remote given that "the vast majority of

---

128. Utility is used here in an economic sense to describe the usefulness of the transaction/outcome of the transaction.

129. *See* Pascoe Pleasence et al., *Wrong About Rights: Public Knowledge of Key Areas of Consumer, Housing and Employment Law in England and Wales*, 80 MOD. L. REV. 836, 839 (2017) (for a broad discussion of influences); *see also* ROBERT C. ELLICKSON, ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES 115 (1991) (exploring how social norms may displace formal sources of law); Janice F. Dyer & Conner Bailey, *A Place to Call Home: Cultural Understandings of Heir Property among Rural African Americans*, 73 RURAL SOC. 317, 317–18 (2008) (for an example of community and cultural norms displacing formal legal expectations in regards to "heir property" amongst African American communities).

130. Winnifred R. Louis & Donald M. Taylor, *Rights and Duties as Group Norms: Implications of Intergroup Research for the Study of Rights and Responsibilities*, *in* THE PSYCHOLOGY OF RIGHTS AND DUTIES: EMPIRICAL CONTRIBUTIONS AND NORMATIVE COMMENTARIES 105, 107 (Norman J. Finkel & Fathali M. Moghaddam eds., 2005).

the population systematically mispredicts . . . the content of the law"[131] for reasons that have been attributed to issue of salience,[132] rationalism (the "need to know"),[133] or ignorance.[134]

Any gap between perceptions of the rights and responsibilities an agreement between parties is presumed to generate, and the rights and responsibilities it does generate (or would likely generate were the agreement to be analyzed judicially) is exacerbated by the fact that the law does not only exist to give effect to private agreements between parties, but also to set broader standards of behavior with which private agreements should conform. In the context of private legal transactions (whether that transaction amounts to an exchange of value, the cessation of a legal relationship through divorce, or the resolution of a dispute) where the parties have the ability to buy knowledge in the form of legal services, it becomes easier to close the gap. However, although this positions a party to better evaluate the risks and benefits associated with agreeing to be bound by certain rights/responsibilities, and to offset or accommodate these realities in the bargain, it also increases transaction costs and inversely diminishes utility. The failure to exhaustively investigate the potential implications of a bargain operates to increase transaction risk but reduce transaction cost, whilst conversely thoroughness reduces risk but increases cost.

Improvements in the storage and processing capacity of computers have enabled the accumulation of vast amounts of data drawn from business, social networks, transaction records, and communications. This information is valuable insofar as it is capable of being translated into strategic insight that informs an assessment of the utility of a bargain at an economic rate. Of course, as the scale of data has expanded, so has the cost of its review. In the context of private ordering this is where AI has principally been used—to improve knowledge of transaction risk (including the transaction risk associated with pursuing litigation) without increasing transaction costs. The application of AI in this context has, as a result, focused

131. Sean Hannon Williams, *Sticky Expectations: Responses to Persistent Over-Optimism in Marriage, Employment Contracts, and Credit Card Use*, 84 NOTRE DAME L. REV. 733, 734 (2009).

132. *See, e.g.*, LaVell E. Saunders, *Collective Ignorance: Public Knowledge of Family Law*, 24 FAMILY COORDINATOR, Jan. 1975, at 69; Jo Casebourne et al., EMPLOYMENT RIGHTS AT WORK: A SURVEY OF EMPLOYEES 2005 (Department of Trade and Industry, 2006); Pascoe Pleasence & Nigel J. Balmer, *Ignorance in Bliss: Modeling Knowledge of Rights in Marriage and Cohabitation*, 46 LAW & SOC'Y REV. 2, 297 (2012).

133. Casebourne, *supra* note 136, at 19.

134. *See, e.g.*, Peter Bowal, *A Study Of Lay Knowledge Of Law In Canada*, IND. INT'L & COMP. L. REV. 121 (1999); PASCOE PLEASENCE & NIGEL J. BALMER, HOW PEOPLE RESOLVE 'LEGAL' PROBLEMS (2014); Catrina Denvir et al., *When legal rights are not a reality: do individuals know their rights and how can we tell?*, 35 J. SOC. WELFARE & FAM. L. 139 (2013); Pleasence et al., *supra* note 129, at 837 n.6.

less on modeling legal reasoning, and more on gathering inputs that might inform the legal (and often the economic) reasoning that takes place during bargaining and negotiation—as the use of e-discovery tools makes clear.

In the litigation setting, litigation proceeds only after a detailed discovery process has taken place in which relevant material is shared between parties.[135] This process first involves in-depth document review in which materials are identified as responsive or unresponsive to a disclosure request, and materials identified as responsive are redacted to uphold legal privilege.[136] Perhaps unsurprisingly, the scope of potentially relevant documents has only increased as a result of the digitization of data, creating an audit trail of correspondence, internet search history, phone records, text messages, electronic transfers, file downloads, and so forth. The amount of data now subject to disclosure vastly increases the time required for manual review and therefore the associated transaction costs. So much so that in some instances it may be preferable to settle a case rather than assume the costs of undertaking disclosure (let alone the cost of other elements of case management and representation).[137] In an effort to offset these costs, AI (largely employing ML techniques, including NLP) has been used to develop computer assisted human review (CAHR) and human-aided computer review (HACR) techniques.[138] This has bred a number of commercial e-discovery software packages, examples of which include kCura Relativity, Ringtail, Logikcull, and Thomson Reuters eDiscovery Point.[139] The growth in the number of commercial e-discovery products made available over the last decade has not been founded on advances in AI, but rather on packaging freely available NLP code within intuitive interfaces. Hence, although these systems may be optimized to perform legal review tasks and reliant on specifically

---

135. FED. R. CIV. P. 26.

136. FED. R. CIV. P. 26(b)(1).

137. See Ross Chaffin, *The Growth of Cost-Shifting in Response to the Rising Cost and Importance of Computerized Data in Litigation*, 59 OKLA. L. REV. 115 (2006), for a relevant example.

138. See Christopher Hogan et al., *Human-Aided Computer Cognition for E-Discovery*, *in* PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 194, 194 (2009), for a deciprtion of the distinction between the two.

139. *See About Us*, RELATIVITY, https://www.relativity.com/about-us/ [https://perma.cc/H8Z6-X7AX]; *Early Case Assessment, Investigations and Document Review*, RINGTAIL, https://www.ringtail.com/ringtail-ediscovery-software/early-case-assessment [https://perma.cc/AKB5-T3PU]; *About Logikcull*, LOGIKCULL, https://logikcull.com/company/#about [https://perma.cc/JZB6-SAA6]; *eDiscovery Point*, THOMSON REUTERS, https://legalsolutions.thomsonreuters.com/law-products/solutions/ediscovery-point [https://perma.cc/9BLF-ZJLR].

developed taxonomies and ontologies (though the level of this optimization has been subject to debate[140]), they are not forms of AI unique to law.

Such tools address a particular gap in the market created by the redundancy of previous approaches. For example, the increasing scale of data has rendered filtering documents by keyword searching (the first implementation of document review) insufficient, and modern filtering techniques have started to take into account the context of words.[141] Other implementations (e.g., in the context of email correspondence) have moved to incorporate knowledge about dependencies between topics.[142] But as is the case with models designed to embody legal reasoning processes, with greater model complexity comes reduced transparency. This is an issue of no small consequence given that discovery is governed by a range of legal obligations and consequences.[143] Efforts to discharge those obligations using electronic tools require legal approval and this can lag behind development. Predictive coding for disclosure (i.e. using ML NLP based techniques to frame a disclosure search request, rather than traditional keyword or Boolean logic searching) has been available for over two decades.[144] Yet in the United States, although e-discovery and technology assisted review have been permitted by way of the 2015 changes to the Federal Civil Procedure Act, and judicially approved in *Moore v. Publicis Groupe*,[145] challenge is possible in instances where it can be shown that the process does not produce reliable and/or proportional results.[146]

Similarly, due diligence, contract review, and lease review processes have also benefitted from ML implementations. This is because in the absence of some form of automation, documents are reviewed in a linear, manual fashion, making it difficult to acquire

---

140. Eugene Yang et al., *Effectiveness Results for Popular e-Discovery Algorithms*, *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 261, 261 (2017).

141. Phuoc An Vo et al., *supra* note 110, at 192.

142. *See* Jyothi K. Vinjumur, *Evaluating Expertise and Sample Bias Effects for Privilege Classification in E-Discovery*, *in* PROCEEDINGS OF THE 15TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 119, 120 (2015).

143. *See, e.g.*, *West African Gas Pipeline Company Ltd v. Willbros Global Holdings Inc* [2012] EWHC 396 (TCC), 141 ConLR 151 (England & Wales); Dan H. Willoughby et al., *Sanctions for E-Discovery Violations: By the Numbers*, 60 DUKE L.J. 789, 791 (2010).

144. *See* Michael Aikenhead, *Legal Knowledge Based Systems: Some Observations on the Future*, WEB J. CURRENT LEGAL ISSUES (1995), https://www.bailii.org/uk/other/journals/WebJCLI/1995/issue2/aiken2.html [https://perma.cc/QDM2-H7UV].

145. Moore v. Publicis Groupe, 287 F.R.D. 182, 183 (S.D.N.Y. Feb. 24, 2012).

146. In England and Wales these methods were not judicially approved as discharging the obligation to perform a "reasonable search" under rule 31.7 of Practice Direction 31B on the disclosure of electronic documents until the 2016 case of *Pyrrho Investments*. *See* Pyrrho Investments Ltd v. MWB Property Ltd & Ors, [2016] EWHC 256 (Ch) (England & Wales), 31B PD (DISCLOSURE OF ELECTRONIC DOCUMENTS).

global appreciation of overall risks. However "entity recognition" software packages designed to extract key details from a vast range of material (structured and unstructured) can vastly accelerate the process of document review.[147] For example, in the context of a business merger requiring the review of vast numbers of mortgage/loan/lease agreements for which a purchaser will assume responsibility, the extraction of key details (such as value, duration, expiry dates, parties, addresses, asset return/yield/encumbrance) enables the quantification of risk and an appreciation of the economic and legal liabilities being assumed by a purchaser, as well as the potential investment yield. These forms of complex analytics bring a range of advantages in assessing the risk/benefits of a transaction, including the accuracy of asset pricing and the discovery of potential legal problems (e.g., through the identification of unusual contract clauses). Whilst these implementations have tended to accompany traditional forms of bargaining (face to face, verbal, written), there have also been efforts to automate elements of the bargaining process, including that of offer and acceptance.

2. *Simulating Bargaining*

Attempts to model the way in which bargains are struck in private (between individual rather than corporate actors), have drawn on theories from the social, behavioral, and economic sciences. More recently, game theory has been used to inform formal and normative models of bargaining[148] and to augment/underpin the dynamics of hybrid rule-based/case-based reasoning software created to simulate bargaining in family law disputes.[149] Though the intent is to simulate "real life" bargaining, the extent to which game theory accurately represents how bargaining agents operate in the real world and the factors that inform their evaluation of the risk/benefit/utility of a bargain, remains subject to doubt. Models based on game theory often assume that agents are perfectly rational and operate to maximize their

---

147. *See iManage RAVN*, IMANAGE, https://imanage.com/product/ravn/ [https://perma.cc/27W2-KBAG]; *About Kira Systems*, KIRA, https://www.kirasystems.com/about/ [https://perma.cc/SKH5-BJWD]; *Homepage*, BRAINSPACE, https://www.brainspace.com/ [https://perma.cc/N7LQ-BJV3]; *About Drooms*, DROOMS, https://drooms.com/en/about [https://perma.cc/HP5H-QW5P].

148. *See, e.g.*, Elisa Burato & Matteo Cristani, *Contract Clause Negotiation by Game Theory*, *in* PROCEEDINGS OF THE 11TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 71 (2007) (discussing the use of game theory to simulate the process of contract clause negotiation).

149. *See, e.g.*, John Zeleznikow & Emilia Bellucci, *Family-Winner: Integrating Game Theory and Heuristics to Provide Negotiation Support*, *in* PROCEEDINGS OF THE SIXTEENTH INTERNATIONAL CONFERENCE ON LEGAL KNOWLEDGE BASED SYSTEMS 21, 22 (2003); Emilia Bellucci & John Zeleznikow, *Family-Negotiator: An Intelligent Tool for Supporting Legal Negotiation in Australian Family Law*, *in* PROCEEDINGS OF THE FOURTH CONFERENCE OF THE INTERNATIONAL SOCIETY FOR DECISION SUPPORT SYSTEMS 359 (1997).

utility in all circumstances, despite research observing that humans "often do not actually have the stable utility functions postulated by both rational-actor theory and usual versions of bounded-rationality theory."[150]

Bargaining is also contingent upon the transmission of beliefs, since each agent must believe that a certain decision is capable of benefiting them both in order for a bargain to be struck/negotiation to be successful. Belief transmission is not static but dynamic, evolving in light of the passing of time, the receipt of new information, the network in which exposure to the belief occurs, and the resistance to belief change exhibited by an agent when faced with contradictory information.[151] It is possible to accommodate some of these dynamic features to, for example, incorporate a measure of authority to represent the weighted plurality of beliefs in an agent's local networks as others have done.[152] Yet, the complexity of human decision-making in respect of law, bargaining, and other difficult problems may not be capable of expression via AI. Arguably, the interaction between agent-based models, social identity theory, and nudge economics that combine to produce a bargain in a given situation are unlikely to accommodate distillation into chains of rules (in the case of symbolic approaches) or intelligible mappings between inputs and outputs (in the case of sub-symbolic/probabilistic methodologies). And, in respect of the latter, there are risks associated with bargaining to a "curve."[153]

Certain software such as "Picture it Settled," which employs negotiation move planning and probabilistic evaluation of case outcome, has enjoyed some commercial success.[154] Other tools, such as eBay's dispute resolution platform, which uses a rules-based decision tree to diagnose problems and propose options for resolution, have come to replace human mediated forms of online dispute resolution (ODR).[155] Nevertheless, simulated models of bargaining remain a niche area, though one that continues to capture the imagination of a range of stakeholders who see it as capable of promoting access to justice and accelerating the bargaining and dispute resolution process between

150. *See, e.g.*, Osoba & Davis, *An Artificial Intelligence/Machine Learning Perspective on Social Simulation: New Data and New Challenges*, *supra* note 121, at 10.

151. *See id.* at 17–18 (showing that if the initial beliefs of all agents are specified and the simulation is then executed, the fraction of the agents believing "something" changes over time).

152. *See id.*

153. *See* Stevenson & Wagoner, *supra* note 31, at 1339.

154. *See About*, PICTURE IT SETTLED, http://www.pictureitsettled.com/about-2/ [https://perma.cc/QU6J-9APD].

155. Ayelet Sela, *Can Computers Be Fair? How Automated and Human-Powered Online Dispute Resolution Affect Procedural Justice in Mediation and Arbitration*, 33 OHIO ST. J. ON DISPUT. RESOL. 91, 93 (2018).

lay parties.[156] Indeed, it remains an active area of research for the likes of Facebook,[157] notwithstanding the fact that implementing such tools in practice relies upon greater structural changes to the legal system.

In alternative dispute resolution systems the law is framed around dispute disposal via settlement,[158] necessitating cooperation of both parties, awareness of what is a fair outcome, and awareness of the consequences attached to non-compliance.[159] In the absence of these conditions being met, the "law" that emerges in the process of bargaining becomes merely a reflection of the power-relations/equality of arms between parties. Systems may incentivize cooperation and/or introduce protections to constrain the actions of those who exhibit poor bargaining abilities, though not without a loss of autonomy. Even with such protections, agreements reached can be reneged upon, with the formal legal system offering the only avenue of recourse. And, whilst this is not an issue unique to bargaining that occurs online, there is reason to believe that technology exacerbates a sense of detachment from responsibility,[160] perhaps more so where autonomy is viewed as being impeded.[161]

Although the law has adopted a technological equivalence approach—a belief that laws and rules should be equivalent in online and offline spaces—this does not always correspond with user perceptions. The disinhibiting effect of online communication leads to an online/offline cognitive divide in which the consequences of actions in the online world are not always seen as translating to the offline world.[162] For this reason enforcement remains a key issue in respect to bargaining

156. Among those interested parties are the Civil Justice Council of England and Wales, who have proposed that an online dispute resolution system for low value civil claims incorporating an "automated negotiation" process would enable better administration of justice, though have remained silent with regards to the more critical details regarding feasibility and function. *See* CIVIL JUSTICE COUNCIL ONLINE DISPUTE RESOLUTION ADVISORY GROUP, ONLINE DISPUTE RESOLUTION FOR LOW VALUE CIVIL CLAIMS 3–5 (2015).

157. Mike Lewis et al., *Deal or No Deal? Training AI Bots to Negotiate*, FACEBOOK ENGINEERING (June 14, 2017), https://code.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/ [https://perma.cc/V35N-N8CY]; *see also* Mike Lewis et al., *Deal or No Deal? End-to-End Learning for Negotiation Dialogues*, *in* PROCEEDINGS OF THE 2017 CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING 2443 (2017).

158. Extrapolating Genn's argument that in the context of alternative dispute resolution, the result is access to a "settlement" and not to "justice." As, "[t]he mediator does not make a judgement about the quality of the settlement." Hazel Genn, *What Is Civil Justice For? Reform, ADR, and Access to Justice*, 24 YALE J.L. & HUMAN. 397, 411 (2012).

159. *See generally id.*

160. *See, e.g.*, Jonathan A. Obar & Anne Oeldorf-Hirsch, *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services*, INFO., COMM. & SOC'Y 1, 1–11 (2016) (discussing the tendency of social media users to accept a platform's privacy policy "without accessing, viewing, or reading any part of it").

161. Mary L. Cummings, *Automation and Accountability in Decision Support System Interface Design*, 32 J. TECH. STUD. 23, 23 (2006).

162. Brian Christopher Jones, *The Online/Offline Cognitive Divide: Implications for Law*, 13 SCRIPTED 83, 87–91 (2016); Obar & Oeldorf-Hirsch, *supra* note 160, at 15.

and failure to accommodate dimensions of enforcement in the online space ultimately diminishes the potential contribution automated negotiation systems might make to access to justice. The seeming lack of consideration as to how enforcement might feature as part of an automated AI-enhanced bargaining system is interesting given that separately, an entire field of enforcement and compliance software has developed as part of the "regtech" (regulatory technology) movement.[163]

C. Public Enforcement & Private Compliance

Technology has been applied in a number of ways in public enforcement and private compliance to support the coercive function of law. Public enforcement refers to the work undertaken by public agencies to monitor adherence to rules or access to entitlements. This activity is not limited to criminal justice agencies, but also incorporates the work of government departments authorized to allocate resources and tasked with monitoring this allocation. As with other domains, software development in the regulatory field have involved both symbolic and non-symbolic approaches to AI, yielding tools to: identify welfare fraud,[164] determine supplementary benefit entitlement,[165] identify tax law abuse,[166] and address conflict of laws in legislative drafting.[167] In criminal justice, the U.S. has led the application of statistical techniques to data containing information about reoffending rates and socio-demographic characteristics to inform parole decisions,[168] with other jurisdictions following suit.[169] ML has also been used to model

---

167. *See, e.g.*, THE REGTECH BOOK: THE FINANCIAL TECHNOLOGY HANDBOOK FOR INVESTORS, ENTREPRENEURS AND VISIONARIES IN REGULATION (Janos Barberis et al. eds., 2019).

168. Amie Meers et al., *Lessons Learnt About Digital Transformation and Public Administration: Centrelink's Online Compliance Intervention*, COMMONWEALTH OMBUDSMAN, July 2017, at 4; ADMINISTRATIVE REVIEW COUNCIL, AUTOMATED ASSISTANCE IN ADMINISTRATIVE DECISION MAKING: REPORT TO THE ATTORNEY-GENERAL 15 (2004).

165. *See* T.J.M. Bench-Capon et al., *Logic Programming for Large Scale Applications in Law: A Formalisation of Supplementary Benefit Legislation*, *in* PROCEEDINGS OF THE 1ST INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 190, 191–92 (1987) (discussions relating to DHSS Demonstrator Project at Imperial University).

166. Erik Hemberg et al., *Tax Non-Compliance Detection Using Co-Evolution of Tax Evasion Risk and Audit Likelihood*, *in* PROCEEDINGS OF THE 15TH INTERNATIONAL CONFERENCE ON ARTIFIFICIAL INTELLIGENCE & LAW 79, 79 (2015).

167. Tingting Li et al., *A Model-Based Approach to the Automatic Revision of Secondary Legislation*, *in* PROCEEDINGS OF THE FOURTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 202, 204 (2013).

168. BERNARD E. HARCOURT, AGAINST PREDICTION : PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE 7–38 (2007).

169. For example, England and Wales have recently reported using historical police records to inform custody risk decisions. *See* Chris Baraniuk, *Durham Police AI to Help with Custody Decisions*, BBC NEWS ONLINE (May 10, 2017), http://www.bbc.co.uk/news/technology-39857645 [https://perma.cc/4DPD-KV2Z].

the "future dangerousness" of parolees and inform sentencing decisions,[170] whilst neural networks have been used to extract "entities" ("facts") from police reports to aid in the criminal investigation process.[171] Further variations in the form of "smart city infrastructure" have been designed to: alert enforcement agencies to gunshots fired in real time by using ML to analyze acoustic sensor data,[172] use AI security cameras for the identification of unattended bags in busy areas,[173] and detect crime clusters for the purposes of resource management.[174]

In the realm of private enforcement, self-governance—in which regulators have taken the veracity of data contained within corporate information returns on trust—has become increasingly uncommon following the 2008 global financial crisis.[175] This has seen regulators permit less autonomy and demand more granular information from regulated entities.[176] These increased regulatory duties have acted as a catalyst for the development of AI tools intended to minimize the cost of regulatory compliance for certain industries. This is of particular relevance in finance, where large exposures, liquidity measures, collateral, capital levels, and stress tests must be reported to regulatory agencies, and where discharging obligations in respect of prudential regulations, data reporting, execution of trades, money laundering, financing of terrorism, due diligence (Know your Customer—KYC), and

---

170. *See generally* RICHARD A. BERK, CRIMINAL JUSTICE FORECASTS OF RISK: A MACHINE LEARNING APPROACH (2012).

171. *See, e.g.*, Michael Chau et al., *Extracting Meaningful Entities from Police Narrative Reports*, *in* PROCEEDINGS OF THE 2002 ANNUAL NATIONAL CONFERENCE ON DIGITAL GOVERNMENT RESEARCH 1 (2002).

172. *See ShotSpotter Technology*, SHOTSPOTTER, http://www.shotspotter.com/technology [https://perma.cc/AX2L-WKMA].

173. *See, e.g.*, *Intel AI Developer Program: Unattended Baggage Detection Using Deep Neural networks in Intel Architecture*, INTEL DEVELOPER ZONE (last updated Apr. 16, 2018), https://software.intel.com/en-us/articles/unattended-baggage-detection-using-deep-neural-networks-in-intel-architecture [https://perma.cc/PEN6-ECUW]; *Movidius*, INTEL MOVIDIUS, https://www.movidius.com/ [https://perma.cc/H3JY-BYH3].

174. *See Homepage*, PREDPOL, http://www.predpol.com/ [https://perma.cc/DVX2-LH2L].

175. This is illustrated with reference to the Supranational Basel II Capital adequacy framework (which specifies the amount of cash reserves that must be retained by a financial institution in order to cover market, credit, and operating risks), in respect of which large financial institutions were permitted to use their own risk management models in determining the level of capital to be held. However, this created an incentive to underestimate credit risk so as to minimize the size of compulsory ring-fenced capital. *See, e.g.*, Douglas W. Arner et al., *FinTech, RegTech, and the Reconceptualization of Financial Regulation*, 37 NW. J. INT'L LAW & BUS. 371, 388 (2017); Harald Benink & George Kaufman, *Turmoil Reveals the Inadequacy of Basel II*, FIN. TIMES (Feb. 27, 2008), https://www.ft.com/content/0e8404a2-e54e-11dc-9334-0000779fd2ac [https://perma.cc/V22D-E8JR].

176. Harriet Agnew, *Andy Haldane Seeks Real-Time Global View*, FIN. TIMES (Oct. 31, 2014), https://www.ft.com/content/4a2fbe2e-6053-11e4-88d1-00144feabdc0 [https://perma.cc/2ATG-2BSK].

data protection can be onerous.[177] AI has also been used to enhance business intelligence, with complex models employed to build consumer credit risk algorithms that draw on information about the debtor, and on the historical transaction (default) patterns of consumers.[178] Variations that have made it to the market include tools that employ a combination of rule-based, ML, and NLP approaches to provide onboarding/customer screening, transaction and trade monitoring, alert management and investigation solutions, and regulatory update dashboards.[179]

Increased regulation and reporting requirements are one way of managing risk, though the pressure placed on organizations and institutions to discharge these responsibilities as efficiently as possible, can lead to the principles and values enshrined in other laws being sacrificed in the service of compliance. The use of AI to discharge regulatory duties may offer public benefit where the costs of compliance are not ultimately passed onto consumers, as may ordinarily be the case. However, technology in the regulatory space may operate to prioritize risk mitigation in favor of other goals and values enshrined by law. For example, the use of ML in retail banking to predict the probability of a customer defaulting on a loan can have perverse consequences. In a regulatory environment that dissuades excessive risk-taking, refusing loans and prioritizing assets with a better risk/yield profile may operate to exclude a wide range of customers who report socio-demographic characteristics that have been historically associated with higher rates of default.[180]

Compliance with regulation, particularly where that compliance involves automation, may create perverse incentives to operate in a way that brings into play other risks.[181] Evidently, this is not a consequence of AI, but may be exacerbated by AI depending on the goals to which algorithms are orientated—exemplifying the broader AI Value

177. *See* Financial Stability Board, *Artificial Intelligence and Machine Learning in Financial Services Market Developments and Financial Stability Implications*, FIN. STABILITY BOARD 1, 1–40 (2017) (providing an overview of developments).

178. Bart van Liebergen, *Machine Learning: A Revolution in Risk Management and Compliance?*, CAPCO INST. J. FIN. TRANSFORMATION 60, 61 (2017).

179. *See, e.g.*, *About Us*, MERLON INTELLIGENCE, https://merlonintelligence.com/about/ (last visited Aug. 1, 2018); *About*, FISCALNOTE, https://fiscalnote.com/about/ (last visited Aug. 1, 2018); *Regulatory Change Management for the Modern Financial*, COMPLIANCE.AI, https://www.compliance.ai/who-we-serve [https://perma.cc/98SV-XPDB].

180. *See, e.g.*, Bryce Goodman & Seth Flaxman, *European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"*, AI MAG., Fall 2017, at 50, 53.

181. For example, Friedman and Kraus argue that Basel rules created perverse incentives to invest in the mortgage-backed securities that caused the global financial crisis. *See* JEFFREY FRIEDMAN & WLADIMIR KRAUS, ENGINEERING THE FINANCIAL CRISIS: SYSTEMIC RISK AND THE FAILURE OF REGULATION 2 (2011).

Alignment discussion.[182] It is not clear how regulatory technology might be programed or how it might respond to ethical dilemmas that occur in particular business contexts. Choices inevitably have to be made but doing so can lead to a form of inequality that is magnified and ultimately structurally embedded as a result of its application at arm's length from human decision-makers.[183]

This only serves to reinforce the fact that in the enforcement of law, as with its interpretation and application, the benefits brought by AI are not evenly distributed. AI tools further the interests, agendas, assumptions, and expectations of their creators. Whilst not all interests are to be construed cynically, it is nonetheless important to try to make sense of the risks posed by various AI systems and approaches; to look beyond specific examples so as to consider the framework that emerges at the intersection of AI and law and in whose interests it serves.

D. Representing Law using AI

AI systems necessarily impose a particular interpretation of what the law is and how its constituent features (interpretation, application, and enforcement) should function. As our review of these systems makes clear, technical complexity can operate to obscure a clear view of the implications associated with implementation. Often these complexities are exacerbated in the market where vendors have a vested interest in generating a sense of awe as to the seemingly sentient/superhuman nature of certain products. However, our understanding of legal AI can be vastly simplified if instead of focusing on the detailed technical intricacies of the technology itself, we focus on the model of law that a given AI approach necessarily imposes. If we are cognizant of the model of law being imposed, then we are better positioned to assess whether that model is appropriate in a given situation.

As our review of the development of law machines reveals, law has been variously represented in AI systems as a product of formalist interpretation of source material (rule-based systems), judicial custom (case-based reasoning), sociological bias (data-driven systems), normatively framed bargaining outcomes (argumentation systems), linguistic interpretation (NLP systems), and regulatory self-governance (RegTech). But, although legal AI has done well to represent the various

---

182. *See, e.g.*, *The Value Alignment Problem*, LEVERHULME CTR. FOR FUTURE INTELLIGENCE, http://lcfi.ac.uk/projects/ai-futures-and-responsibility/value-alignment-problem/ [https://perma.cc/JR56-4WY5].

183. *See, e.g.*, Marion Oswald, *Algorithm-Assisted Decision-making in the Public Sector: Framing the Issues Using Administrative Law Rules Governing Discretionary Power*, 376 PHIL. TRANSACTIONS ROYAL SOC'Y. 1, 8 (2018); MEERS ET AL., *supra* note 132, at 2; ALGORITHEMWATCH, AUTOMATING SOCIETY TAKING STOCK OF AUTOMATED DECISION-MAKING IN THE EU 8 (Matthias Spielkamp ed., 2019); Thomas J. Barth & Eddy Arnold, *Artificial Intelligence and Administrative Discretion Implications for Public Administration*, 29 AM. REV. PUB. ADMIN. 332, 346–47 (1999).

manifestations of law and the theoretical models thought to underpin these manifestations, it does less well to integrate these models in a way that recognizes their pluralistic interplay. Though not always the case,[184] each incarnation of legal AI tends to reflect only one interpretation of how we understand "law," and in doing so implies that it is possible for law (as a system) to be represented via a single explanatory model that is mutually exclusive from and superior to other models of understanding.

We oppose this view: as a complex system, the crux of "law" lies not in one model of interpretation, but in the interaction between all of these models of interpretation. Legal systems exist not just to give effect to rights and responsibilities, but also to provide an infrastructure within which those rights and responsibilities can be contested, developed, adapted, and evolved to better balance interests, reflect social needs, enact justice, and accommodate rule of law constraints. As such, it is not possible to account for the many and varied constructions of law using one computational (or for that matter epistemological) paradigm. Any application of AI that adopts one computational representation of law to the exclusion of all others, must consider the impact of doing so. This is well known by many working within the field of AI, though it is a point at risk of being overlooked by implementing organizations in the public and private sector. The model that law takes when enshrined in the infrastructure of technology also has certain consequences for the rule of law. These are issues we discuss further in Part III.

III. THE RULE OF (MACHINE-MADE) LAW

Efforts to measure the strength of the rule of law in a particular jurisdiction have varied in respect to method and focus. In line with the United Nation's approach,[185] the Organisation for Economic Cooperation and Development's (OECD) method of evaluation relies on institutional features that install the conditions necessary for the rule of law to flourish.[186] Accordingly, objective indicators of progress, such as the creation of legal guarantees of due process or a legal framework to guarantee impartiality of the judiciary, are favored over evaluation as

184. *See, e.g.*, Zeleznikow & Stranieri, *supra* note 12, at 185.

185. UNITED NATIONS DEP'T OF PEACEKEEPING OPERATIONS & OFFICE OF THE HIGH COMM'R FOR HUMAN RIGHTS, THE UNITED NATIONS RULE OF LAW INDICATORS: IMPLEMENTATION GUIDE AND PROJECT TOOLS (2011).

186. BERENSCHOT & IMAGOS, THEMATIC EVALUATION OF RULE OF LAW, JUDICIAL REFORM AND FIGHT AGAINST CORRUPTION AND ORGANISED CRIME IN THE WESTERN BALKANS–LOT 3 (2013); OECD, EVALUATION OF GOVERNANCE, RULE OF LAW, JUDICIARY REFORM AND FIGHT AGAINST CORRUPTION AND ORGANISED CRIME IN THE WESTERN BALKANS: LOT 2–FINAL REPORT (2012).

to whether adherence to these legal requirements occurs in practice.[187] This top-down institutional focus does not consider what the citizen-experience of justice might reveal about the state of the rule of law in a jurisdiction. Structural reforms intended to enhance the rule of law speak to intent but not to operability, which can only be understood in the context of the experience of justice.[188]

Increasingly, calls have been made to recognize bottom-up indicators that provide qualitative and experiential insight into how the rule of law is actualized in practice by those subject to the law.[189] For an individual, taking formal action to enforce a right is only likely to merit the effort if the system within which interpretation and enforcement occurs is perceived to be functional, impartial, fair, and accessible. These perceptions color informal dispute resolution, which is said to occur in the "shadow of the law," with parties having reference to the likely interpretation/sanction imposed by the court.[190] By extension, issues are resolved not just in the "shadow of the law," but in the "shadow of the rule of law," with parties considering the functionality, impartiality, fairness, and accessibility of institutions trusted to make decisions. Formal systems governed by weak legal order have a flow on effects reflected in the "private ordering" that occurs in relation to informal dispute resolution.[191]

For AI, both the top-down (structural/institutional/constitutional/formal) and bottom-up (experiential) dimensions of the rule of law are of importance. The potential of AI as a means by which to improve justice must be assessed with reference to the institutional and experiential rule of law impact as well as the distribution of this impact. In

---

187. *See. e.g.*, BERENSCHOT & IMAGOS, *supra* note 190, at 35–36; OECD, *supra* note 190, at 16–20; MARTIN GRAMATIKOV & RONALD JANSE, CONCEPT PAPER: MONITORING AND EVALUATION OF THE RULE OF LAW AND JUSTICE IN THE EU: STATUS QUO AND THE WAY AHEAD? 6–8 (2012).

188. *See* GRAMATIKOV & JANSE, *supra* note 187, at 9.

189. *See, e.g.*, *The World Justice Project: General Population 2016–Opinion poll*, WORLD JUSTICE PROJECT (2016), https://worldjusticeproject.org/sites/default/files/documents/gpp_questionnaire_2016_final.pdf [https://perma.cc/W64L-LTNY] (describing the World Justice Index—an effort to derive a more subjective citizen-centric sociologically based insight by drawing on public perceptions of the likely outcome in different hypothetical scenarios so as to construct an understanding of the prevailing legal order).

190. Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950, 997 (1979).

191. Hazel Genn, *What Is Civil Justice For? Reform, ADR, and Access to Justice*, 24 YALE J.L. & HUMAN. 397, 398 (2012) [hereinafter Genn, *What is Civil Justice For?*]. As Genn explains,

> Authoritative judicial determination has a critical public function in common-law systems, creating the framework or the 'shadow' in which the settlement of disputes can be achieved. That it is underpinned by the coercive power of the state provides the background threat that brings unwilling litigants to the negotiating table and makes it possible for weaker parties to enforce their rights and to expose wrongdoing.

Id.

the section that follows we consider the key rule of law considerations that are raised in the application of AI to law. We look specifically at the capacity of AI to balance procedural versus substantive justice, and to enhance (or undermine) the neutrality, transparency, and accessibility of the legal system, and the autonomy of legal agents.

A. *Substantive versus Procedural Justice*

A key structural dimension of the rule of the law is the way in which legal outcomes balance procedural versus substantive justice. Whilst procedural justice is enshrined in the rules that govern how the law is made and applied, substantive justice is bound in the consequences that derive from this process of "making" or "applying." These goals are diametrically opposed. To permit a higher degree of substantive justice is to permit a lower degree of procedural justice.[192] Whilst for some, (notably Jerome Frank) substantive justice is viewed as taking priority such that where justice and the rule of law diverge; "the rule of law is pernicious to the extent that it detracts from achieving justice."[193] For others, such as Robert H. Bork, there is no universal agreement as to what constitutes substantive justice other than by "reference to some system of moral or ethical values that has no objective or intrinsic validity of its own and about which men can and do differ."[194] On this view, adherence to the procedural constraints of the rule of law is the closest to what we might call "objective" justice.[195]

If, as has been argued, procedure is more important than outcome, then rule-based AI systems offer a means by which to safeguard consistency of legal decision-making and provide a level of certainty in the law that does not exist at present. Such systems would operate to emphasize the scientific character of the law in which key dimensions of the rule of law, namely conformity to reason, uniformity, and certitude are prioritized above all else.[196] Indeed, there are those who advocate rule-based reasoning as a means of enhancing the rule of law, observing the scientific (formalistic) character of law as the only path by which to safeguard full, equal, and exact justice.[197] Rule-based judgment systems may have the effect of instigating a shift away from what D'Amato describes as living under the rule of persons, as opposed to living under the rule of law.[198] But in evaluating the merits

---

192. *See* Barnett, *supra* note 20, at 598–99.

193. *Id.* at 597–98 (paraphrasing Jerome Frank).

194. Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1, 10 (1971).

195. Barnett, *supra* note 20, at 598.

196. *But see* Marcin Matczak, *Why Judicial Formalism is Incompatible with the Rule of Law*, 31 CANADIAN J.L. & JURIS. 61, 63 (2018).

197. Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605, 605 (1908).

198. Bork, *supra* note 198, at 10.

of this approach, we must consider whether all that is lost when we remove the humanness of legal judgment is the institutionalization of arbitrariness, discretion, and/or bias.[199]

The impact of this loss turns on whether the ends (fair outcomes) are seen to occupy a position of greater importance than the means (consistent procedure). Although the flexibility that we come to associate with human judgment generates a degree of uncertainty and ambiguity, it may also fulfill a positive function in the application and interpretation of law. Firstly, because it enables "just" outcomes to be reached, and secondly, because legal evolution relies on this discretion, allowing legal interpretation to act in reflection of and reaction to broader social change.[200] The architecture of rule-based reasoning systems is such that all the information must be contained within the system from the start. Such systems cannot respond to the evolution of the law without requiring frequent rebuilding, nor would evolution of the law be permitted if formal systems were to replace the existing legal ecology. There is reason to be cautious of making the law into a system of rules in the service of technology,[201] such that we might concede, as Weizenbaum has done, that although computers can make judicial decisions, they ought not be given such tasks.[202] Crucially, despite their ability to learn, it is not clear that machine-learning tools escape these issues of legal stagnancy, and in fact, when developed for the purpose of legal decision-making, may only serve to exacerbate stagnancy by rendering new outcomes a facsimile of those that have gone before.

B. Institutionalizing Bias

As a tool to assist legal decision makers, the predictions derived from data-driven systems offer insight into factors that influence an outcome. But whilst they may operate as a form of system monitoring, capable of flagging up administrative or judicial decisions that suspiciously diverge from previous decisions, they also present a degree of risk. Although ML methods are not numerically stagnant, as Bayesian approaches enable "rules" (as defined by the numerical relationships or weights between inputs and outputs) to update upon receipt of new data (in the form of new case decisions[203]), ML systems can be

---

199. Anthony D'Amato, *Can/Should Computers Replace Judges?*, 11 GA. L. REV. 1277, 1279 (1977).

200. *See, e.g.*, *id.* at 1281; Pound, *supra* note 197, at 606.

201. Carl F. Stover, *Technology and Law—A Look Ahead*, 4 M.U.L.L. MOD. USES LOGIC L. 1, 4–5 (1963).

202. POPPLE, *supra* note 78, at 10, 52 n.42 (citing JOSEPH WEIZENBAUM, COMPUTER POWER AND HUMAN REASON: FROM JUDGMENT TO CALCULATION 226–27 (1976)).

203. *See generally* DAVID BARBER, BAYESIAN REASONING AND MACHINE LEARNING (2012).

functionally stagnant, representing a self-fulfilling prophecy of sorts.[204] Realists argue that legal interpretation is the product of a whole range of different factors, only some of which are rooted in deductive or analogical reasoning.[205] Producing a ML model that determines the outcome of a case by replicating the influence of certain factors in existing data acts only to propagate similar such outcomes.[206]

In other words, data can encode biases due to the human decisions that this data represents, and ML reliant on this data can operate to institutionalize this bias.[207] Whilst claims of "biased" computer programs suggest an ethical failure on behalf of designers, such problems may more commonly reflect methodological and technical issues with data, only some of which might be avoidable. So, whilst ML has the potential to eliminate bias from decision-making, allowing for models to be tweaked by reducing the effect of certain irrelevant characteristics on an outcome,[208] this assumes that the underlying data is not compromised (biased) through improper sampling, collection, or simply because those using it have failed to appreciate the systemic or historic inequalities it reflects.[209]

The likelihood of institutionalizing bias within a model bias is not so remote. Widely used criminal risk prediction software in the U.S.

---

204. *See generally* Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values, Washington, D.C.*, (May 2014), https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf [https://perma.cc/9V8B-ZCGR].

205. *See generally* Pound, *Theories of Law*, *supra* note 20, at 114–16; Leiter, *Legal Realism*, *supra* note 20, at 261.

206. *See, e.g.*, Executive Office of the President, *supra* note 208; Bennett-Moses & Chan, *supra note* 96, at 648; Centre for Data Ethics and Innovation (CDEI) & Cabinet Office, *Interim Report: Review into Bias in Algorithmic Decision-Making* (last updated July 25, 2019), https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making [https://perma.cc/5TZF-WB5X].

207. *See generally* Chris DeBrusk, *The Risk of Machine-Learning Bias (and How to Prevent It)*, MIT SLOAN MGMT. REV. (Mar. 26, 2018), https://sloanreview.mit.edu/article/the-risk-of-machine-learning-bias-and-how-to-prevent-it [https://perma.cc/TF59-GNU8]; Will Knight, *Forget Killer Robots—Bias is the Real AI Danger*, MIT TECH. REV. (Oct. 3, 2017), https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/ [https://perma.cc/PP5V-VPJW]; Hannah Devlin, *AI Programs Exhibit Racial and Gender Biases, Research Reveals*, GUARDIAN (Apr. 13, 2017), https://www.theguardian.com/technology/2017/apr/13/ai-programs-exhibit-racist-and-sexist-biases-research-reveals [https://perma.cc/LYY6-PXC3]; OSONDE OSOBA & WILLIAM WELSER IV, AN INTELLIGENCE IN OUR IMAGE: THE RISKS OF BIAS AND ERRORS IN ARTIFICIAL INTELLIGENCE (2017).

208. *See, e.g.*, Sam Corbett-Davies et al., *Algorithmic Decision Making and the Cost of Fairness*, *in* PROCEEDINGS OF THE 23RD ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 797, 805 (2017) (on the use of constrained optimization techniques to impose algorithmic fairness).

209. Monika Ermert, *WEF Davos: Who Will Own the Knowledge Produced from "Our" Data by Machines?*, INTELL. PROP. WATCH (Jan. 18, 2017), https://www.ip-watch.org/2017/01/18/wef-davos-will-knowledge-produced-data-machines/ [https://perma.cc/Q98F-EV6Z].

has already been shown to routinely assign higher levels of risk to the assessments of black subjects.[210] This is also true in respect of NLP, where the gendered nature of language has ramifications for the models produced from natural language data.[211] While the detection and removal of bias from text is an active area of research,[212] the issue of latent bias in NLP is far from solved. Similar such challenges arise in the application of AI to the fields of finance and medicine, where there is the potential for AI agents to inadvertently discriminate against a loan applicant or a medical patient on the basis of age, gender, sexuality, etc.[213] Moreover, in crime detection and enforcement, the implications arising from the imposition of certain technologies may run counter to the goal the technology is intended to achieve. Predictive policing software that identifies crime hot-spots risks reinforcing a vicious cycle of increased police presence in neighborhoods already subject to over-policing.[214] These flaws can be understood if we recognize the method being used to develop the software, but they are not necessarily straightforward to pinpoint where the complexity of a model hinders its explainability, or the commercial interests of a developer inhibits disclosure and restricts transparency.[215]

---

210. Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/NC5S-L79Y].

211. Tolga Bolukbasi et al., *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*, *in* 30TH CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS 4356 (2016); Aylin Caliskan et al., *Semantics Derived Automatically from Language Corpora Contain Human-Like Biases*, 356 SCIENCE 183, 184–85 (2017); Su Lin Blodgett & Brendan O'Connor, *Racial Disparity in Natural Language Processing: A Case Study of Social Media African-American English*, *in* FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY IN MACHINE LEARNING (FAT/ML) WORKSHOP (2017); Rachael Tatman, *Gender and Dialect Bias in YouTube's Automatic Captions*, *in* PROCEEDINGS OF THE FIRST WORKSHOP ON ETHICS IN NATURAL LANGUAGE PROCESSING 53, 57 (2017).

212. *See e.g.*, Marta Recasens et al., *Linguistic Models for Analyzing and Detecting Biased Language*, *in* PROCEEDINGS OF THE 51ST ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (VOLUME 1: LONG PAPERS) 1650 (2013); Liye Fu et al., *Tie-Breaker: Using Language Models to Quantify Gender Bias in Sports Journalism*, *in* PROCEEDINGS OF THE IJCAI WORKSHOP ON NLP MEETS JOURNALISM 1 (2016).

213. Danton S. Char et al., *Implementing Machine Learning in Health Care–Addressing Ethical Challenges*, 378 NEW ENGL. J. MED. 981, 981–83 (2018) (discussing such inadvertent discrimination in the medical field).

214. *See, e.g.*, P. Jeffrey Brantingham et al., *Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial*, 5 STAT. & PUB. POL'Y 1, 2–6 (2018); Danielle Ensign et al., *Runaway Feedback Loops in Predictive Policing*, *in* 81 PROCEEDINGS OF MACHINE LEARNING RESEARCH: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 1, 2 (2018).

215. Kate Crawford, *Artificial Intelligence's White Guy Problem*, N.Y. TIMES (June 25, 2016), https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html [https://perma.cc/58NX-AKKV]; Katherine Freeman, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, 18 N.C. J.L. & TECH. 75 (2016); Ellora Israni, *Algorithmic Due Process: Mistaken Accountability and Attribution in State v. Loomis*, JOLT DIGEST (Aug. 31 2017),

C. Transparency

The transparency of legal decision-making plays a crucial role in shaping public perceptions of the fairness of the justice system. It speaks to the perceived structural and experiential integrity of the legal system and operates as a barometer for the relative strength of the rule of law. AI, particularly ML, raises concerns regarding the explainability and robustness of the decisions made. Automated decision-making systems implemented with the intent of facilitating access to justice by enhancing consistency may in practice operate as a barrier to the obtainment of fair outcomes. This risk can arise as a function of design or of implementation, with the latter risk magnified where the process by which automated decisions are reached are procedurally and substantively different from those processes adhered to by human decision-makers.[216]

Thus, there is particular concern associated with applying deep learning in certain contexts (including law) where explainability is key, as the outcomes produced cannot always be clearly and systematically interpreted.[217] It may be difficult for a clinician to explain to a patient why they have been given a prognosis of five years to live rather than ten as expected, just as it may be difficult for a loan officer to justify the denial of a loan to a prospective customer without a substantive basis for doing so. Similarly, a decision to find a defendant guilty, or to sentence someone to more than the average incarceration period for a particular crime, needs to be supported with clear reasons. The use of such algorithms in a public or private administrative capacity (e.g., to determine access to legal aid, to determine access to the entitlements of citizenship, to vet potential employees, or to assess an applicant's credit risk), involves interaction with the law and with characteristics protected by law. These are the sorts of protections that can be easily (deliberately or inadvertently) subverted by machine decision-making, and ascribing responsibility for failure may prove difficult.

There are also constraints imposed as a result of technological infrastructures. For example, accuracy in ML is based on the raw

---

https://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and-attribution-in-state-v-loomis-1 [https://perma.cc/VFF8-QZ39].

216. As shown in an inquiry into the Australian Department of Human Services (DHS) "Online Compliance Intervention" (OCI) system, launched in July 2016 to detect discrepancies between the income welfare benefit clients reported to DHS and the income they reported to the Australian Tax Office (ATO). This revealed that whilst the algorithm employed to identify inconsistencies replicated the calculations of human decision makers, the information that informed a calculation was qualitatively different. *See* AMIE MEERS ET AL., *supra* note 164, at 4.

217. NUFFIELD COUNCIL ON BIOETHICS, BIOETHICS BRIEFING NOTE: AI IN HEALTHCARE AND RESEARCH 1, 4 (May 2018), http://nuffieldbioethics.org/project/briefing-notes/artificial-intelligence-ai-healthcare-research [https://perma.cc/4GAQ-8KNH].

prediction performance of an algorithm exposed to "unseen data" (or new cases/combinations of facts) as compared to accuracy achieved in other similar such studies.[218] It is not always possible to achieve 100% accuracy, and some level of inaccuracy is always accepted. In instances where an algorithm is charged with determining whether a picture is of a cat or a dog, the ramifications of an error (at least in the abstract) are relatively minor. However, the impact of error is of far greater magnitude in regard to legal decision-making, particularly where systems fulfill a "gatekeeping" function. Whilst a right to explanation exists in the U.S. in relation to credit score,[219] it is not clear the extent to which the right exists in relation to other forms of administrative decision-making.[220] Certainly if the court's decision in *State v. Loomis* is indicative, that right might extend only to ensuring that decision-makers (as opposed to subjects) are provided with a written warning as to the dangers of reliance on certain algorithmic assessments.[221] These issues force us to consider the level of machine error we are willing to accept given that this machine error compounds any existing human error in the labeled data on which an algorithm is trained.

D. Access to Justice

Access to justice is taken to mean access to the institutions of justice, such as courts, dispute resolution services, administrative appeal mechanisms to legal services, access to knowledge about one's rights, entitlements and obligations, and the ability (personal or structural) to exercise those rights and uphold those responsibilities.[222] Open

---

218. Foster Provost et al., *The Case Against Accuracy Estimation for Comparing Induction Algorithms*, *in* PROCEEDINGS OF THE FIFTEENTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 445–453 (1998). This is particularly common when research is carried out on well-known tasks on publicly available datasets (as often happens in the field of computer vision, for example). *See, e.g.*, Richard Dinga et al., *Beyond Accuracy: Measures For Assessing Machine Learning Models, Pitfalls and Guidelines***,** BIORXIV (Aug. 22, 2019), biorxiv.org/content/10.1101/743138v1 [https://perma.cc/YJ4G-XDDJ].

219. Equal Credit Opportunity Act, 12 U.S.C. §§ 5512, 5581 (2012).

220. Similarly, although the General Data Protection Regulations offer some safeguards to EU residents by restricting automated decision-making and creating a right to explanation, the range of limitations and exceptions that apply have been roundly critiqued. *See* Goodman & Flaxman, *supra* note 180, at 55; Maja Brkan, *AI-Supported Decision-Making under the General Data Protection Regulation*, *in* 617 PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 3, 5 (2017).

221. 881 N.W.2d 749, 767 (Wis. 2016) (deciding whether failure to reveal the basis of proprietary risk assessment software used in sentencing violated the defendant's due process rights).

222. Whilst acknowledging that "in practice access to justice defies definition" and is "used as a handle to justify all sorts of policies designed to have quite different outcomes," Hazel Genn provides the following definition:

> At its most basic it is about access to procedures for making rights effective through state-sponsored public and fair dispute resolution processes. It implies

access to the courts is of critical importance—courts operate to uphold standards of behavior, which in turn frame the informal application, interpretation, and enforcement of law.[223] Courts are the only means by which to legitimize the threat of enforcement and the view that real or perceived consequences may actually arise.[224] Yet, whilst access to (formal and informal) adjudication systems are important, mere access is not sufficient if would-be users lack the knowledge and capacity to make use of this access.[225] In practice, access requires that the public have the requisite knowledge, understanding, confidence, and capability to pursue or defend a legal claim—with or without professional assistance.

To date, much has been said about the capacity for AI tools to aid in facilitating access to justice; either through increasing public knowledge of the law, democratizing access to the legal services market by reducing costs, and/or increasing the capacity of adjudicative institutions to handle claims by making processes more efficient.[226] Indeed, a number of examples previously provided suggest progress in these areas. Further, as-yet unrealized ambitions to use NLP techniques to translate (user) natural language to the legalese required in a "particulars of claim" statement, to fill out a court document, or to automate points of contest in a negotiation, represent a natural evolution of much of the AI and law work already done. Yet, despite much talk about the potential utility of AI systems in providing

---

> equal access to authoritative enforceable rulings and outcomes that reflect the merits of the case in light of relevant legal principles. It does not imply that laws are necessarily just, but that individuals have a fair opportunity for their rights to be determined according to the prevailing promulgated rules.

HAZEL GENN, JUDGING CIVIL JUSTICE 115 (2010) [hereinafter GENN, JUDGING CIVIL JUSTICE]. *See also* Rebecca L. Sandefur, *Access to What?*, 148 DAEDALUS 49, 50 (2019) (More recently, Sandefur has proposed "[t]here is access when disputes and problems governed by civil law . . . resolve with results that satisfy legal norms."); DEBORAH L. RHODE, ACCESS TO JUSTICE (2004).

223. *See* Mnookin & Kornhauser, *supra* note 190, at 990–92.

224. As Genn explains,

> Authoritative judicial determination has a critical public function in common-law systems, creating the framework or the "shadow" in which the settlement of disputes can be achieved. That it is underpinned by the coercive power of the state provides the background threat that brings unwilling litigants to the negotiating table and makes it possible for weaker parties to enforce their rights and to expose wrongdoing.

Genn, *What is Civil Justice For?*, *supra* note 158, at 398. Genn further explains, "While the reality is that most cases settle, a flow of adjudicated cases is necessary to provide guidance on the law and, most importantly, to create the credible threat of litigation if settlement is not achieved." GENN, JUDGING CIVIL JUSTICE, *supra* note 222, at 21.

225. Rt. Hon Lord Bingham of Cornhill, *supra* note 20, at 6.

226. *See, e.g.*, ONLINE DISPUTE RESOLUTION ADVISORY GROUP, *supra* note 127, at 3; RICHARD SUSSKIND & DANIEL SUSSKIND, THE FUTURE OF THE PROFESSIONS: HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS (2015).

advice and support in the civil justice resolution process, there are no obvious examples where these ambitions have been realized.[227]

The "trickle down" benefits of AI in democratizing access to legal services for individual consumers of modest means, has been minimal at best. Indeed, in some cases AI may actively (though not always obviously) undermine efforts to widen access to justice. Tools that predict the likely success of the case may diminish a lawyer's willingness to represent a client in court, no matter how important the cause. Such information may also influence the likelihood of a plaintiff's ability to secure third party financing—a phenomenon that has emerged over the last decade and is particularly prominent in class action and tort cases.[228] Rule-based or data-driven systems implemented as a gateway to accessing a court (with or without representation) also generate further cause for concern. These technologies represent efficiency gains for certain actors (usually law firms or government agencies) rather than mechanisms by which justice may be made more accessible to the public.[229] As a result, they call into question the effect AI may have in distorting equality of arms.

The use of AI in law also raises a series of existential questions regarding the purpose of law. If AI affects a shift in the locus of legal decision-making from humans to technology, what role is there for law to remain the "primary instrument to guide and sustain legitimate expectations between those who share jurisdiction"?[230] Where AI is charged with moderating interactions between actors (customers, clients, citizens) and agents (businesses, government, organizations), technology may displace the function of the law in setting legitimate expectations to guide shadow bargaining. In such instances, bargaining no longer takes place in the shadow of the law but in the shadow of automated decision-making infrastructures. By extension, success may have less to do with legal merit and more to do with the likelihood

227. For example, in Sandefur's review of 322 access to justice technologies intended for non-lawyers, only two purported to use AI technology, with a further one indicating an intention to draw on AI in the development of the software. REBECCA L. SANDEFUR, LEGAL TECH FOR NON-LAWYERS: REPORT OF THE SURVEY OF US TECHNOLOGIES 31, 53, 57 (2019).

228. Jason Krause, *Third-party financing is growing, and lawyers are big players*, ABA J. (July 1, 2016), http://www.abajournal.com/magazine/%0Barticle/third_party_financing_is_growing_and_lawyers_are_big_players [https://perma.cc/EWP9-CKEQ].

229. For example, the related example produced by the UK government's introduction of a telephone gateway for access to legal aid in 2012. Although not AI enhanced, the gateway was variously accused of creating a hurdle to access to justice. Ben Hickman & David Oldfield, *Keys to the Gateway: An Independent Review of the Mandatory Civil Legal Advice Gateway*, PUB. L. PROJECT (2015), https://publiclawproject.org.uk/wp-content/uploads/data/resources/199/Keys-to-the-Gateway-An-Independent-Review-of-the-Mandatory-CLA-Gateway.pdf [https://perma.cc/5JXS-VWXY].

230. Mireille Hildebrandt, *Law as Information in the Era of Data-Driven Agency*, 79 MOD. L. REV. 1, 8 (2016).

of overturning an automated decision-making process.[231] Such issues, particularly where they involve inequality of arms between parties and where they relate to "settlement" transactions contingent on an evaluation of utility and risk/benefit, are likely to undermine bargaining power. This has the effect of displacing the procedural function of the rule of law that exists to enable people to stand up for their rights[232] and the due process protections enshrined in the U.S. Constitution. More perverse are applications where Big Data is used to gain advantage by exploiting weaknesses in the prevailing legal order. For example, in Article 33 of the Justice Reform Act France recently banned using Big Data about judges to exploit systemic bias and gain an advantage over adversaries.[233]

Where AI systems extend to bargaining processes, bargaining becomes oriented around arriving at an outcome rather than a just outcome[234] and further, runs the risk of equating the value of bargaining/negotiation with the value of the agreement reached. Such tools might increase the likelihood of parties reaching an agreement by ensuring a solution in all instances in which parties exhibit overlapping bargaining ranges, but at what cost for procedural justice or perceptions of fairness? Individuals are said to care deeply about the process by which an outcome is reached,[235] and automated systems are likely to not only render these processes opaque, but to subsequently diminish belief in the fairness of the steps by which a decision has been made. Such tools, whether deployed in or outside of

231. *See* Lilian Edwards & Michael Veale, *Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?*, 16 IEEE SECURITY & PRIVACY 46, 51–52 (2018) (discussing the "right to an explanation" in the context of data protection and subject access requests).

232. *Id*. at 2.

233. Loi 2019–222 du 23 mars 2019 de programmation 2018–2022 et de réforme pour la justice [Law 2019-222 of Mar. 23, 2019 on the Reform of Justice], JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE [J.O.] [OFFICIAL GAZETTE OF FRANCE], Mar. 21, 2019, Article 33 (stating that "[t]he identity data of magistrates and members of the Registry cannot be reused with the purpose or effect of evaluating, analyzing, comparing or predicting their actual or alleged professional practices.").

234. This parallels the "anti-adjudication and anti-law discourse" implicit in policy reforms occurring in the UK from the 1980s onwards, which has seen access to justice reframed as access to an outcome and which has instigated a shift away from access to formal legal structures towards mediation and alternative dispute resolution. Genn, *What is Civil Justice For?*, *supra* note 191, at 409. As a result, Genn argues that access to justice has been redefined not as access to the courts or to a "just settlement" but as access to a "settlement"—as demonstrated by the fact that mediators do not make a judgment about the quality of the settlement. *Id.* at 411. The same is true of bargaining systems, which are not focused on matters of fairness, but rather on expedience of settlement. *See, e.g.*, Julia Hörnle, *Encouraging Online Dispute Resolution in the EU and Beyond*, 38 EUR. L. REV 187, 208 (2013) (taking issue with systems which prioritize efficiency, cost-effectiveness, and automation at the expense of due process).

239. Rebecca Hollander-Blumoff & Tom R. Tyler, *Procedural Justice in Negotiation: Procedural Fairness, Outcome Acceptance, and Integrative Potential*, 33 L. & SOC. INQUIRY 473, 477 (2008).

the negotiation context, might also be seen to encroach on legal agency given that the exercise of justice and access to it becomes bound by the parameters set by technology designers and/or the limits of a given technology. As Pound has noted, "The effect of all system is apt to be petrifaction of the subject systematized."[236]

E. Legal Agency and the Distribution of Benefits

Legal agents have been described as "participant in a distinctly constituted social condition" within which they are constituted as "responsible persons" capable of engaging in "purposive action . . . understand[ing] and follow[ing] practical standards," and "accountable for the decisions they make with respect to what the law requires of them."[237] Those who design and implement legal AI systems do so in order to further a particular agenda, undermining the aspiration that the law might exist as a neutral mechanism balancing competing interests in society. AI systems have the potential to diminish agency by requiring that bargaining/application/adjudication occurs in conformance with the constraints of a particular system or process. When the choice to make a decision is removed or when decisions amount to nothing more than a fait accompli, in which no space is given to revision, reflection, or debate, then personal agency is necessarily constrained. Behavior mediated via data-driven systems with their own form of agency is not independently exercised.[238]

Law is not merely information about the legal effect of one's behaviors, but an agent that gives substance to that effect. The idea of prospectivity as a fundamental feature of the rule of law implies that those subject to the law should be capable of anticipating the legal effect of their actions. Yet as Hildebrand observes, data-driven techniques are used more often to anticipate and pre-empt the behavior of legal subjects and as such these tools inform the likelihood of taking an action and not the implications of doing so.[239] That AI tools may operate to constrain autonomy and limit agency is of fundamental concern given that the creation, interpretation, application, and enforcement of law draws legitimacy from the presumed autonomy of the actors involved. In a rule of law context, these issues speak to

236. Pound, *supra* note 197, at 608.

237. Stefano Bertea, *Legal Form and Agency: Variations on Two Central Themes in Fuller's Legal Theory*, 5 JURIS. 96, 98 (2014).

238. MIREILLE HILDEBRANDT, SMART TECHNOLOGIES AND THE END(S) OF LAW: NOVEL ENTANGLEMENTS OF LAW AND TECHNOLOGY 8–9 (2015).

239. *Id.*; Hildebrandt, *supra* note 230, at 10.

where accountability might vest for mistakes or poor judgment, as well as highlighting those for whom the rewards of AI are likely to accumulate.[240]

Asking who is likely to benefit from the implementation of AI in different settings and what this benefit entails, is key to understanding the potential implications of that implementation on legal agency and outcome. In medicine, the lines appear more clear-cut. Benefit is defined as that which prolongs life and minimizes suffering. To this end, AI has been used to model the underlying causes of diseases,[241] churn through biomedical research papers and flag compounds which may be cures to a given disease,[242] classify specific genes according to their role in disease development,[243] and analyze biomolecular structure.[244] It has also been used in diagnostic settings in conjunction with radiology and ultrasound,[245] to detect brain tumors,[246] quantify risk of Alzheimer's,[247] diagnose liver diseases,[248] recognize malignant prostate tissue,[249] analyze retinal scans,[250] and construct models of the heart to predict survival chances for patients with pulmonary

240. Extending legal personhood to AI has heralded calls for further exploration of the legal status of autonomous agents and clarification regarding the status, rights, and obligations of "electronic personalities." *See, e.g.*, Argyro Karanasiou & Dimitris Pinotsis, *Towards a Legal Definition of Machine Intelligence: The Argument for Artificial Personhood in the Age of Deep Learning*, *in* PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 119 (2017).

241. *See, e.g.*, *Publications*, BERG, https://berghealth.com/publications/ [https://perma.cc/S73Q-XNRM].

242. *See* Ken Mulvany, *Developing Pharmaceuticals Efficiently with Artificial Intelligence*, INT'L BUS. TIMES (Feb. 28, 2018), https://www.ibtimes.co.uk/developing-pharmaceuticals-efficiently-artificial-intelligence-1664431 [https://perma.cc/6N3Y-8575].

243. *See* WuXi NextCODE, *Artificial Intelligence Used to Advance Precision Therapy for Rare Genetic Disorders of Obesity*, PR NEWSWIRE (Jan. 9, 2018), https://www.prnewswire.com/news-releases/wuxi-nextcode-artificial-intelligence-used-to-advance-precision-therapy-for-rare-genetic-disorders-of-obesity-300579919.html [https://perma.cc/RB44-SXK2].

244. *See* Nic Fleming, *Computer-Calculated Compounds*, 557 NATURE 55, 57 (2018).

245. *See* SANKETH VEDULA ET AL., TOWARDS CT-QUALITY ULTRASOUND IMAGING USING DEEP LEARNING (2017).

250. *See* Guotai Wang et al., *Automatic Brain Tumor Segmentation using Cascaded Anisotropic Convolutional Neural Networks*, *in* MICCAI BRAINLESION WORKSHOP 178 (2017).

247. Ammarah Farooq et al., *A Deep CNN Based Multi-Class Classification of Alzheimer's Disease Using MRI*, *in* 2017 IEEE INTERNATIONAL CONFERENCE ON IMAGING SYSTEMS AND TECHNIQUES 1, 2 (2017).

248. *See* Koichi Ogawa et al., *Computer-Aided Diagnostic System for Diffuse Liver Diseases with Ultrasonography by Neural Networks*, 45 IEEE TRANSACTIONS ON NUCLEAR SCI. 3069, 3069 (1998).

249. *See* Tillman Loch et al., *Artificial Neural Network Analysis (ANNA) of Prostatic Transrectal Ultrasound*, 39 PROSTATE 198, 200 (1999).

250. *See* Jeffrey De Fauw et al., *Automated Analysis of Retinal Imaging Using Machine Learning Techniques for Computer Vision*, 5 F1000RESEARCH 1573 (2016).

hypertension.[251] On the face of it, these implementations appear to serve the interests of the public, but a closer assessment might call into question the disruptive effect of these tools on existing ecologies. In medicine as in law, the advice giving/treatment process relies upon a functioning lawyer-client/doctor-patient relationship. A third-party actor in the form of an ML model may complicate this interpersonal connection and weaken the strength of communication. As with most technological innovation, some stakeholders are likely to benefit from the implementation of AI in certain contexts more than others.[252]

AI in law has been driven by those with the most to gain from its introduction. This has implicated corporate law firms who have mobilized AI tools to increase efficiency, protect profit margins, and increase the ratio of support staff, relative to qualified legal staff. This concentration of activity by corporate stakeholders reflects stark investment realities as to who has the capital and incentive to invest. These pragmatic constraints only serve to reinforce how the benefits of AI may be unequally distributed.[253] In the public sector, this imbalance may see enhanced efficiency at the cost of diminished fairness, with the interests of government (spend reduction) prioritized over the interests of justice.[254] Elsewhere, the imbalance may enable a third

255. *See* Timothy J.W. Dawes et al., *Machine Learning of Three-dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study*, 283 RADIOLOGY 381, 382 (2017).

252. In respect of medicine the response of the Royal College of General Practitioners to Babylon Health's publication is informative. *See* Royal College of General Practitioners, *Apps and Algorithms May "Support but will Never Replace" GPs, Says RCGP*, RCGP NEWS (June 27, 2018), https://www.rcgp.org.uk/about-us/news/2018/june/apps-and-algorithms-may-support-but-will-never-replace-gps-says-rcgp.aspx [https://perma.cc/7VVV-52JV] (suggesting that Babylon's service "cherry-picks" more straightforward patients, due to the bias in those who can access the app).

253. In some cases, the "access to justice" angle appears to be a detour on the road to commercialisation. In 2016 software start-up LawGeex invited the public to upload their employment contracts so that software could scan the contracts so as to identify unusual contract terms or conditions, in what was framed as an access to justice offer. *See LawGeex Now Reviews Employment Contracts Free, LawGeex* (Apr. 19, 2016), https://web.archive.org/web/20160807140250/http://blog.lawgeex.com:80/lawgeex-now-reviews-employment-contracts-free/ [https://perma.cc/6FCM-KJL8] ('The more contracts that we review, the better our machine learning algorithms get, and the more people we can help."). The offer was short-lived, as their website now returns a 409 error and with no evidence that this personal option is still available. *See Not Found, Error 404*, LAWGEEX, http://blog.lawgeex.com/lawgeex-now-reviews-employment-contracts-free/ [https://perma.cc/75WZ-DHA7].

254. As demonstrated by the Australian Department of Human Services (DHS) "Online Compliance Intervention" (OCI) system launched in July 2016. MEERS ET AL., *supra* note 183, at 1. The system was designed to automate the investigation and debt raising process where DHS detected a discrepancy between the amount of Pay-As-You-Go (PAYG) income reported to DHS (in respect of welfare benefit receipt) and the amount of PAYG income reported to the Australian Tax Office (ATO). *Id.* Intended to increase the capacity for discrepancy identification and investigation beyond the level achievable via human oversight, the system generated an exponential increase in the number of discrepancies identified. *Id.* Sustained public criticism emerged as to the transparency, fairness, and usability of the system

party to benefit at the expense of others. Whilst the AI enhanced-eBay dispute resolution system is routinely lauded, it is worth noting that most disputes end in negative feedback being left, followed by termination of the bargain.[255] This negates most of the economic benefits of the transaction whilst increasing the cost of bargaining by way of inconvenience/postage costs/transaction fees incurred.

Although AI and the increased automation it facilitates may reduce bias, corruption, inefficiency, inconsistency, and inaccessibility, the efficacy of a particular technology remains contingent on the broader eco-system within which it is deployed and to whom it is made accessible. As Hildebrand proposes "algorithms to delve and further develop legal knowledge should primarily serve those subject to the law, not first and foremost those administering the law"[256] and should be available to all, rather than "restricted to those willing and able to pay the fees of the corporate law firms that have the capital to invest."[257]

F. Assessing Risk versus Benefit

For those who perceive the law as a reflection of the personal views of those people in charge of enacting, applying, and enforcing it, legal rules are "objects of discourse, not objects with a concrete nature."[258] In this context, formalizing the law in a rule-based system, classifying future cases with reference to historic cases, linguistic corpora, or psychometric data, silences this discourse by removing the space required for it to thrive. In a constitutional democracy, law sustains the balance of power between citizens, businesses, and the state, providing a

---

and the vulnerable user-group subject to OCI debt notifications. *Id.* at 2. The resulting Senate inquiry revealed that whilst the algorithmic calculation employed to identify inconsistencies replicated that of human decision makers, the information that informed a calculation was qualitatively different. *Id.* Human investigators were required to attempt "every possible means of obtaining the actual income information." *Id.* at 3. Under OCI the onus for providing information rested on the claimant, and where that onus was not discharged within a specified timeframe, averaged ATO income was used to calculate debt owing. *Id.* This averaging was necessary in order to achieve the underlying efficiency objective the tool was intended to introduce. *Id.* at 4.

> To enable it to automate debt raising in situations where earnings information was not forthcoming from the customer, DHS decided to accept the best *already available* evidence to calculate an approximate debt figure by averaging ATO data, rather than using its information gathering powers to obtain verified fortnightly data to calculate an exact debt figure. This decision was fundamental to the efficiency and scale of the system, because it meant that compliance officers did not have to manually intervene to obtain fortnightly payroll data.

Id. at 4.

255. Lilian Edwards & Ashley Theunissen, *Creating Trust and Satisfaction Online: How Important is ADR? The UK eBay Experience*, *in* 21ST BILETA CONFERENCE: GLOBALISATION AND HARMONISATION IN TECHNOLOGY LAW 1, 16 (2006).

256. Hildebrandt, *supra* note 230, at 10.

257. *Id.* at 11.

258. Leith, *Fundamental Errors in Legal Logic Programming*, *supra* note 45, at 548.

structure within which agents in society can interact. The fact that technology also has the power to regulate these interactions implies that "legal and technological instruments are not exchangeable tools to achieve specific policy objectives, depending on which tool is more efficient or effective."[259] Adopting such an attitude puts us at risk of a system in which technology displaces the regulatory power of the law—governing the law itself, rather than being governed by it.[260]

This necessarily implies a conservative approach to using AI technologies as an instrument by which to safeguard fundamental principles of the rule of law. It also urges caution in assuming that such technologies can improve the institutional and structural integrity of legal systems, without the need for further (human) checks and balances. This is particularly relevant given that AI models of law are necessarily constrained by the limits of what is technologically possible at the point of development. Whilst ordinarily the law operates to constrain technology, any effort to enshrine the "law" in an AI system necessarily allows this technology to constrain the interpretation, application, and enforcement of the law. The law cannot be seen to exist as separate from the apparatus that gives it effect. Where that apparatus is technology, the subject of embodiment ("the law") becomes defined by reference to what the object of embodiment (the technological architecture) makes possible.

Recognizing this, and the model of law imposed by a particular AI system, is of value in determining the impact that might result from implementation. However, introspection tends to follow rather than precede implementation past the point at which the influence of that technology has already taken hold. Who has not questioned the validity of a colored squiggly line appearing under a sentence in MS Word, yet deferred to the system in preference to continuing to work under an accusation of grammatical ineptitude? This pervasive influence leads McGee and Eriksson to refer to MS Word as "the invisible grammarian"[261] and reinforces the prescience of Weiser's 1991 claim that "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it."[262] It is only upon reflection we can start to see what might have been possible were we more alive to developments as they occurred.

---

259. Mireille Hildebrandt, *A Vision of Ambient Law*, *in* REGULATING TECHNOLOGIES: LEGAL FUTURES, REGULATORY FRAMES AND TECHNOLOGICAL FIXES 175, 178 (Roger Brownsword & Karen Yeung eds., 2008).

260. *Id.* at 178–80.

261. Tim McGee & Patricia Ericsson, *The Politics Of The Program: MS WORD as the Invisible Grammarian*, 19 COMPUTERS & COMPOSITION 453, 466 (2002).

262. Mark Weiser, *The Computer for the 21st Century*, 265 SCI. AM. 94, 94 (1991).

Our analysis of the rule of law impact of AI has necessarily adopted a critical tone, looking more forcefully at the potential negative rather than positive effects. That is not to say that AI cannot make a positive contribution to the functioning of the legal system. Human decision makers are not infallible, and there are instances where AI can be used to bolster, rather than undermine, the integrity and operation of the rule of law. It is certainly feasible that AI could play a role in remedying defects in the rule of law—democratizing access to knowledge about rights and the mechanisms by which to enforce these rights. Tools employed in the public sector may play a role in making the process of justice more efficient, and in levelling the playing field rather than distorting it. Certainly, is it not impossible to think the vast amounts of investment capital flowing into legal technology might yield dividends for all of society? Nevertheless, we must consider the underlying objectives to which such tools are directed and the implications attached to the representation of law being adopted in any given instance. There are a number of pragmatic obstacles that may dictate an agenda of AI development that serves private interests over and above public interests. These obstacles, discussed further in Part IV, provide some insight into the developments we might expect to see in the coming years, and who is likely to be setting the agenda for progress.

IV. AGENDAS & OBSTACLES

Discussions as to the role of AI in law tend to occupy two different ends of the spectrum. Optimists lean towards controversial assertions that technology will displace lawyers and legal expertise because it is becoming more advanced year-on-year.[263] Whilst the skeptics advance the view that technology will never be able to undertake the work of a lawyer because a lawyer's work involves a complex interplay of different skills.[264] Those who argue for the former generally overestimate the speed of progress and underestimate the challenges. Those who argue for the latter tend to presume that the job of the lawyer remains stagnant, or alternatively, perceive incremental change as stagnancy.[265] At least for the next decade, if not for the foreseeable future, pragmatic constraints related to: (A) the commercial incentives, (B) access to data, and (C) the skills gap, will see the reality of legal

263. *See, e.g.*, RICHARD SUSSKIND, THE FUTURE OF LAW: FACING THE CHALLENGES OF INFORMATION TECHNOLOGY (2d ed. 1998); RICHARD SUSSKIND & DANIEL SUSSKIND, THE FUTURE OF THE PROFESSIONS, *supra* note 188, at 68–69.

264. *See, e.g.*, Dana Remus & Frank Levy, Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law (Nov. 27, 2016) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2701092 [https://perma.cc/FWD3-2QJG].

265. Change blindness explains this tendency to overlook changes that would normally be easily observed. Daniel J. Simons & Ronald A. Rensink, *Change Blindness: Past, Present, and Future*, 9 TRENDS COGNITIVE SCI. 16, 19 (2005).

technology development fall between these two extremes. In the following section we consider each of these pragmatic constraints in turn.

A. Commercial Incentives

As earlier discussions have revealed, the finance and banking industry has been seen as an early adopter of AI, particularly when compared to legal services. This is perhaps not so surprising when we consider the numerical nature of most of the data involved in finance, the financial resources available when computing power remained an expensive commodity, the industry's access to highly specialized skill sets (and the ability to pay high salaries to attract talent), and the strong financial incentives that exist to price assets properly (for example, options and other derivatives) and predict them into the future. Early adoption in finance is even less surprising given that ML represents a natural extension of many of the common technical analyses conducted by systematic traders in investment banks and hedge funds.[266]

By contrast, the technology procurement decisions of law firms have typically been driven by a desire to avoid the loss of clients or limit exposure to liability.[267] Law is a knowledge-based occupation and there has been resistance to the idea that this knowledge can or should be commoditized through the use of technology.[268] Moreover, hourly-billing models, only recently eroded by changes to the market, have arguably disincentivized the pursuit of process efficiencies.[269] Yet, in line with the trends exhibited by their clients, there is growing expectation that law firms (particularly those in the service of large financial institutions) will utilize technology. As such, for most organizations who shop around for legal services on the basis of price, the competitiveness of a quote for legal services will be strongly linked to the number of man hours versus computer hours required.

In spite of claims of slow uptake, the way in which technology has been adopted by law firms does not demarcate the legal industry as uniquely antediluvian. Adoption or pursuit of new technological advancements across the legal industry have not been homogenous. Legal service providers have exhibited evidence of engagement at

---

266. For example, Chen and Lian's Protrader expert system was able to predict the 87 point drop in Dow Jones Industrial Average in 1986. K.C. Chen & Ting-peng Liang, *PROTRADER: An Expert System for Program Trading*, 15 MANAGERIAL FIN. 1, 1 (1989).

267. Legal Technology Insider, *No IT Please, We're Making Enough Money Already!*, LEGAL TECH. INSIDER (Apr. 1, 2004), https://www.legaltechnology.com/wp-content/uploads/2012/01/lti160.pdf [https://perma.cc/LMJ3-K5U6].

268. Maurits Barendrecht, *Legal Aid, Accessible Courts or Legal Information? Three Access to Justice Strategies Compared*, 2011 GLOBAL JURIST Issue 1, Art. 6, at 12; RICHARD SUSSKIND, THE END OF LAWYERS? RETHINKING THE NATURE OF LEGAL SERVICES 3 (2008).

269. Oskamp & Lauritsen, *supra* note 2, at 232.

both the incubation and deployment stages, and the timing and extent of engagement may have less to do with protectionism than is often suggested.[270] Dickerson, for example, has advanced the view that larger firms (more than a dozen lawyers) tend to disproportionately reap the benefits of technology adoption, and this might help explain why these entities appear to be leading the charge in respect of technology research and development.[271] The adoption or development of new tools also requires heavy internal investment and an appetite for risk. These requirements do not necessarily reflect the organizational reality of law and the demand for consensus that the partnership model begets. So, whilst some legal service providers have invested in software spinouts, this is a strategy that represents a clear departure from the core business of legal service delivery.[272]

Doing so also introduces additional risk, since the benefits of technology adoption are not always clear from the outset. Technology adoption requires the support of (public, private, and corporate) clients, and ML can raise particular challenges in implementation. Data protection requirements and the sensitivities of corporate clients, particularly financial institutions, means that cloud storage of client data is often prohibited.[273] ML technologies must be hosted on-site, a solution that vendors (particularly small start-ups) cannot always offer. Data stored by firms in externally hosted clouds remove a level of control, and whilst encryption technologies minimize third party threats, data remains accessible to a cloud host.

Where data does reside in the cloud, there is a further concern regarding intellectual property. ML systems learn from an individual's interaction with the system and where a cloud-based solution is used, an individual is interacting with the software provider's systems and servers.[274] Not all welcome the fact that software developers benefit from the expertise of professional users. These intellectual property considerations also constrain the practicality of joint ventures. Where cooperation occurs, it does so more often in respect of products that are

---

270. SUSSKIND & SUSSKIND, THE FUTURE OF THE PROFESSIONS, *supra* note 188, at 67–68.

271. F. Reed Dickerson, *Electronic Computers and the Practical Lawyer*, 14 J. LEGAL EDUC. 485, 487 (1961).

272. *See, e.g.*, Bruce S. Tether, *Who Co-Operates for Innovation, and Why? An Empirical Analysis*, 31 RES. POL'Y 947 (2000) (discussing the incentives and disincentives associated with technology, partnerships, and innovation).

273. *See, e.g.*, Neil Hodge & Ravi Meah, *Don't Sleepwalk into the Cloud – The Challenges for Law Firms and their Clients*, LEGAL WEEK (June 25, 2015), https://www.law.com/legalweek/2015/06/25/dont-sleepwalk-into-the-cloud-the-challenges-for-law-firms-and-their-clients-2/ [https://perma.cc/8AYZ-6UQU]. Although, there appears to be an emerging uptake of cloud technologies within the profession. *See, e.g.*, ILTA & INSIDELEGAL, 2016 ILTA/INSIDELEGAL TECHNOLOGY PURCHASING SURVEY (2016). Notably, cloud technologies bring data protection and ethics obligations.

274. For a broad discussion of this IP issue, see Ermert, supra note 209.

new to the market, rather than new to a firm, with firms tending to collaborate with non-competitors to develop products with mass market, rather than firm level appeal.[275] This removes what might otherwise be seen as an incompatible tension between a firm's interest in safeguarding a tool where it bestows a competitive advantage, and a developer's interest in selling this tool to a firm's competitors.

A focus on mass-market products resolves the intellectual property and competition challenge, yet simultaneously constrains the eco-system for innovation. It suggests an agenda of future development led by large law firms in the service of large clients and growth of AI in RegTech and corporate law fields rather than access to justice. The latter being a field where the enthusiasm around AI-based technologies wax and wane over shorter timescales than longer-term public sector decision-making can accommodate and where data is not always available in the form (or at the scale) required.

B. Access to Data

The predictive ability of an ML system remains substantially more dependent on the features selected and data quality/quantity used, than the algorithm (or statistical model) employed. In certain circumstances, data will be readily available and provided in a format that lends itself to analysis. Large-scale datasets of U.S. Supreme Court cases yield a rich source of information prime for analysis and for the development of related analytical tools. Other publicly available large-scale data sets (for example, Enron discovery documents) provide source material by which to develop and refine new legal software tools, whilst privately held datasets give rise to collaborations between start-up developers and law firms.[276] Yet whilst digitization has made certain forms of data abundant, it is not always the sort of data that is required for the analysis of law.

In contrast to the U.S., which provides Public Access to Court Electronic Records (PACER) for the network of federal courts, open publishing requirements are often patchy in other jurisdictions. So, although the U.S. PACER system has come under attack for the high

275. *Id.*

276. For example, Ross Intelligence utilized Baker Hostetler's 27 terabyte's of data pertaining to the Bernard Madoff case to help learn bankruptcy law. Ross Intelligence, *ROSS Intelligence Announces Partnership with BakerHostetler*, PRNEWSWIRE (May 5, 2016), https://www.prnewswire.com/news-releases/ross-intelligence-announces-partnership-with-bakerhostetler-300264039.html [https://perma.cc/35DC-57KF]. It should be noted that much of the publicity surrounding Ross Intelligence reported that the firm was "hiring" technology, implying that the system came ready-made. *Id.* In a reality, Ross Intelligence approached Baker Hostetler at an early stage in development knowing that the firm had a large amount of bankruptcy data, with a view to having the firm trial the product to develop the system's understanding and knowledge of bankruptcy law. *Id.*

costs of access and a range of other inadequacies,[277] proponents of the open data movement in the U.S. have increasingly made inroads into democratizing access in a way not achieved elsewhere.[278] In England and Wales for example, there exists a highly fragmented data environment in which the publication of legal decisions has been privatized since inception, with "law reporters and publishers tak[ing] the view that the copy of the judge's text which they hold is their intellectual property." [279] As such, the work of bringing together and providing free and open access to legal judgments has been left to charitable organizations, such as the British and Irish Legal Information Institute (BAILII).[280] Even taking into account resources such as PACER in the U.S., BAILII in the U.K., or the Australian equivalent (AUSTLII),[281] we observe less effort directed towards preserving or making accessible a whole range of potentially useful data collected by public agencies, including e-bundles and administrative data.

The availability of public data is just one dimension of the data challenge. Law firms may hold a whole range of business intelligence, including contracts, legal briefs, legal research, emails, and/or correspondence that can be mined for insights. However, this data is often held over a number of different jurisdictions (each with their own data protection laws), a number of different systems (current and legacy), and firms may impose different data storage systems across different practice areas or locations.[282] From this emerges a series of unanswered questions regarding the extent to which emerging AI technologies will be able to handle siloed data in a way that integrates it more efficiently than the current data pipeline process demands.[283]

It is not just the availability of data that is important, but also the form it takes. The use of ML on non-numerical data, such as natural

---

277. Jeff Roberts, *Why the Federal Court Record System PACER is so Broken, and How to Fix it*, GIGADOM (Aug. 27, 2014), https://gigaom.com/2014/08/27/why-the-federal-court-record-system-pacer-is-so-broken-and-how-to-fix-it/ [https://perma.cc/T749-DNHR].

278. Stevenson & Wagoner, *supra* note 31, at 1361.

279. Philip Leith & Cynthia Fellows, *Enabling Free On-line Access to UK Law Reports: The Copyright Problem*, 18 INT'L J.L. & INFO. TECH 72, 80 (2010).

280. *British and Irish Legal Information Institute*, BAILII, https://www.bailii.org/ [https://perma.cc/BBH6-2SPR].

285. *About*, AUSTRALIAN LEGAL INFORMATION INSTITUTE, https://www.austlii.edu.au/about.html [https://perma.cc/JF9K-QE52].

282. *See, e.g.*, *The Future Architecture of Law Firm Information: A New Foundation for Information Exchange*, HUBBARD ONE & THOMSON REUTERS, https://legalsupportnetwork.co.uk/sites/default/files/HubbardOne_Whitepaper_OneViewFINAL_0811.pdf [https://perma.cc/N6CQ-SFR2].

283. *See, e.g.*, Ignacio Terrizzano et al., *Data Wrangling: The Challenging Journey from the Wild to the Lake*, *in* 7TH BIENNIAL CONFERENCE ON INNOVATIVE DATA SYSTEMS RESEARCH (CIDR '15) 4, 4–5 (2015); Alon Halevy et al., *Managing Google's Data Lake: An Overview of the GOODS System*, 39 BULL. TECHNICAL COMMITTEE ON DATA ENGINEERING 5–14 (2016).

language, requires both an understanding of linguistics as well as massive amounts of training data to infer relationships between entities. It is not difficult to tag a word as a verb or a phrase as passive, but higher-level relationships that are inherent in understanding law require exponentially larger volumes of appropriately (i.e., human) annotated data as well as very sophisticated NLP algorithms.[284] For this reason, associating sentiment with certain word phrases will require fewer training examples as compared to ingesting legalese, suggesting appropriate prior cases, or summarizing an internal narrative and establishing a causal relationship. Most NLP systems remain largely context-specific, and do not generalize well to bodies of text that differ in nature to the training data—both subject matter and temporal considerations are relevant.[285] It has previously been observed that because many NLP tools are trained on data that is now more than 20 years old, they exhibit superior performance for text written by older users.[286] This is of particular relevance to law, given that frequent comparison is made between historical and modern texts.

These challenges belie some of the limitations seen in previous studies and discussed at various points in the preceding sections, notably: the development of models based on very small training sets,[287]

---

284. Ramon F. Astudillo et al., *Learning Word Representations from Scarce and Noisy Data with Embedding Sub-spaces*, *in* PROCEEDINGS OF THE 53RD ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS AND THE 7TH INTERNATIONAL JOINT CONFERENCE ON NATURAL LANGUAGE PROCESSING 1074, 1074 (2015) ("The success of supervised systems largely depends on the amount and quality of the available training data, oftentimes, even more than the particular choice of learning algorithm (Banko and Brill, 2001). Labeled data is, however, expensive to obtain, while unlabeled data is widely available"). *See also* Tomas Mikolov et al., *Efficient Estimation of Word Representations in Vector Space*, CORNELL U. ARXIV (2013), https://arxiv.org/pdf/1301.3781.pdf [https://perma.cc/2VZH-9JQ8]; Kirk Roberts, *Assessing the Corpus Size vs. Similarity Trade-off for Word Embeddings in Clinical NLP*, *in* PROCEEDINGS OF THE CLINICAL NATURAL LANGUAGE PROCESSING WORKSHOP 54–63 (2016).

285. *See, e.g.*, Allyson Ettinger et al., *Towards Linguistically Generalizable NLP Systems: A Workshop and Shared Task*, *in* PROCEEDINGS OF THE FIRST WORKSHOP ON BUILDING LINGUISTICALLY GENERALIZABLE NLP SYSTEMS 1 (2017); Ana Marasović, *NLP's generalization problem, and how researchers are tackling it*, GRADIENT (Aug. 22, 2018), https://thegradient.pub/frontiers-of-generalization-in-natural-language-processing/ [https://perma.cc/GR32-VJB8].

286. Dirk Hovy & Anders Søgaard, *Tagging Performance Correlates with Author Age*, *in* PROCEEDINGS OF THE 53RD ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS AND THE 7TH INTERNATIONAL JOINT CONFERENCE ON NATURAL LANGUAGE PROCESSING 483, 487 (2015).

287. *See, e.g.*, Bochereau et al., *supra* note 102, at 90. This paper uses what would be considered in contemporary research an extremely small number of examples (378) when training their model to discriminate between annulled and confirmed cases, particularly in the context of the number of inputs they use. *Id.* Though this may reflect the lack of digitized data available for training such algorithms in the early 1990s, as well as to a lesser extent the computational power and *ipso facto* the time it would have taken to train the neural network they employ.

the use of hypothetical data to simulate real cases,[288] and the use of data that reflects only part of the phenomenon under observation.[289] Court data poses a particular challenge, not only because published decisions are difficult to acquire, but also because any attempt to predict certain outcomes, such as court decisions, requires more than just data from the court of interest. Cases that reach the upper courts are distinct in nature. This poses fewer problems if the intention of a statistical, ML, or data mining task is simply to look at cases that reach a particular higher court. However, if the intention is to understand which cases reach a court and how they differ from cases that do not progress as far, then the dataset must include cases from lower courts.[290] This means any model produced using this data set is heavily conditioned on the fact that an applicant has progressed this far, rendering such findings less relevant for people at earlier stages of the process.[291] The same issues arise with respect to precedent versus common cases. Law relies on precedent, but this is reflected in the mass of common cases, the vast majority of which remain unreported.

In circumstances where an appetite for investment and the necessary data is available, a further challenge presents itself in terms of the availability of the expertise needed to bring ideas to fruition; an issue we refer to as "the skills gap."

C. The Skills Gap

Expertise at the intersection of law and technology is a relative rarity.[292] In the majority of jurisdictions, lack of exposure to opportunities to acquire quantitative and technological skills during legal study

---

288. For an example of use, see Zeleznikow & Stranieri, *supra* note 12, at 186; for commentary of use, see Hunter, *supra* note 41, at 59–60.

289. This is the case with any algorithms attempting to predict which cases end up in court where the only data used is data drawn from case files. In these instances, the absence of data relating to those cases that do not end up at court is problematic. See, for example, the impact of the absence of full data in Naomi Burstyner et al., *Why Do Some Civil Cases End Up in A Full Hearing? Formulating Litigation and Process Referral Indicia Through Text Analysis*, 25 J. JUD. ADMIN. 257–95 (2016); *see also* Chen & Eagel, *supra* note 88, at 238.

290. In their work, Chen and Eagel explicitly identify this challenge by acknowledging that they were only able to investigate data relating to applicants who have made it to the refugee court system. Chen & Eagel, *supra* note 88, at 237. Another important issue that Chen and Eagel explicitly note is how to deal with missing data, with 80% of the cases in their dataset missing at least one feature. *Id.* at 238.

291. Capturing the experiences of those who do not make it to court typically require the use of wide-scale, expensively compiled legal need survey data. *See* Charles E. Clark & Emma Corstvet, *The Lawyer and the Public: An A.A.L.S. Survey*, 47 YALE L.J. 1272 (1938) (the first of these surveys); PASCOE PLEASENCE ET AL., PATHS TO JUSTICE: A PAST, PRESENT AND FUTURE ROADMAP (2013) (summary of current surveys and impact).

292. Catrina Denvir, *Scaling the Gap: Legal Education and Data Literacy*, *in* MODERNISING LEGAL EDUCATION 73 (Catrina Denvir ed., 2020).

culminates in legal AI being relegated to a fringe activity in law schools.[293] Whilst it may not be necessary to have legal, computing, *and* quantitative skills in order to make a contribution to research and development in legal AI, the early domination of the field by logic programmers (residing ostensibly in computer science departments) was not without criticism. Not because it was impossible for a logician to understand law, but because it was observed that logic programmers all too often failed to move beyond their epistemological view of the world as a computational model resulting in the development of logic models that did not accord with how law operated in practice.[294]

The skills gap presents implications for development of the field, as well validating broader concerns regarding the potential rule of law impact and appropriateness of the models developed. As data mining interfaces begin to democratize access to ML techniques, rendering software more accessible to those without technical expertise, the ethical risks of data-driven technologies increase. Having data is only one dimension to solving a complex problem. The inability to explain patterns can actually be exacerbated when applying ML techniques, because the process of analysis is divorced from the traditional scientific method. In law, traditional qualitative analysis provides a conceptual underpinning that aids quantitative analysis. The data does not do all the work for us, and what it reveals is only useful (and only transferrable to other domains) if it is anchored in a broader theory or hypothesis about how certain phenomena operate or interrelate.[295] This disconnect between "measurable data streams and meaningful explanatory theories to frame the data" is defined as the "theory-data gap," it represents "a key barrier to meaningful social and behavioral modeling," demanding a "move from purely data-driven work to theory-informed work" and a need to "tighten the iterative loop between theory and data analysis."[296] It is a problem common not just to law, but also to medical applications of AI.[297] Moreover, this is not just an issue in respect of ML-based reasoning systems. As Ashley noted in respect of analogical reasoning "analogy is a way of stating

293. For a more detailed exposition, see Philip Leith, *IT and Law, and Law Schools*, 14 INT'L. REV. L. COMPUTERS & TECH. 171 (2000).

294. Leith, *Fundamental Errors in Legal Logic Programming*, *supra* note 45, at 552.

299. *See, e.g.*, Peter Norvig, *All We Want are the Facts, Ma'am*, NORVIG.COM, http://norvig.com/fact-check.html [https://perma.cc/FQ3U-ZSGL].

296. OSOBA & DAVIS, *supra* note 100, at iii.

297. Whilst methods of statistical, data-centric modeling were not unheard of during the early stage of medical AI development. *See, e.g.*, J.R. Staniland et al., *Clinical Presentation of Acute Abdomen: Study of 600 Patients*, 3 BRITISH MED. J. 393, 394 (1972). They were not widely applied in part because the biomedical understanding thought necessary to properly model the underlying processes was lacking. SZOLOVITS, *supra* note 59, at 6.

a conclusion, not reaching one, and theory must do the real work, where theory is the principle that links cases or that separates them."[298]

New variations of ML do not address issues of skills shortage, nor do they do away with the need for an underlying theoretical conceptualization of the construct that a ML model is attempting to represent. NLP offers an instrumental case study in this regard. Previous epochs of NLP have been characterized by great collaboration between domain specialists (linguists) and computer scientists/AI researchers. The Cocke-Younger-Kamasi algorithm [299] recognizes whether a sentence is consistent with a given set of grammatical rules, and generates a chart (parse tree) to reflect the structure of the sentence. [300] Notably, this algorithm is informed by grammatical theory developed by Chomsky[301] as well as efficient computational techniques developed by Bellman,[302] and is thus a good example of the interplay between linguistics and computer science. Admittedly, computer scientists with little linguistic understanding can design deep-learning algorithms: a word-embedding algorithm can learn the quantitative representation of the text on which a deep neural network can be trained. However, throwing data at a problem is only possible with access to huge training sets (of the type retained by large technology companies) and more limited training sets are likely to result in over-fitting and/or increased error/uncertainty. Rather than giving rise to opportunities for innovation, these pragmatic constraints implicitly and explicitly shape the role that AI occupies in and outside of law, and at times, actively militate against innovation.

D. The Shape of AI to Come

It is possible for law to be reflected by AI systems in a variety of different ways, and continued technological development is likely to yield further methods by which to model the law and legal reasoning. Whether or not these systems are designed, developed, and deployed to benefit the majority or the minority remains to be seen. Legal AI

---

302. Kevin Ashley et al., *Legal Reasoning and Artificial Intelligence: How Computers "Think" Like Lawyers*, 8 U. CHI. L. SCH. ROUNDTABLE 2, 20 (2001).

299. This algorithm was derived by three researchers independently. *See* Daniel H. Younger, *Recognition and Parsing of Context-Free Languages in Time n3*, 10 INFO. & CONTROL 189, 189–90 (1967); T. KASAMI, AN EFFICIENT RECOGNITION AND SYNTAX-ANALYSIS ALGORITHM FOR CONTEXT-FREE LANGUAGES (1966); JOHN COCKE & J.T. SCHWARTZ, PROGRAMMING LANGUAGES AND THEIR COMPILERS: PRELIMINARY NOTES (1969).

300. Younger, *supra* note 303, at 189–90; KASAMI, *supra* note 303; COCKE & SCHWARTZ, *supra* note 303.

301. Noam Chomsky, *On Certain Formal Properties of Grammars*, 2 INFO. & CONTROL 137 (1959).

302. Such as dynamic programming: Richard Ernest Bellman, *The Theory of Dynamic Programming*, 60 BULL. AM. MATHEMATICAL SOC'Y 503 (1954).

does not (in and of itself) signal the beginning of a new era of tools capable of making good on the promise of universal access to justice or unfailing adherence to the rule of law. As our analysis of the pragmatic challenges reveal, these objectives cannot be realized without a supporting infrastructure. Nevertheless, given the exponential advances in data and processing power seen over the last two decades, we are still at a relatively early stage of data-driven AI development. That there are few examples to illustrate how the application of AI to law and the representation of law via AI might enhance the rule of law and access to justice is not to say that such uses are not possible. Nor is it the case that in order for one group to gain another group must necessarily lose, even if the conditions steering development in the field tend to favor private over public interests.

However, safeguarding the rule of law and the protections enshrined in the constitution requires a commitment to using AI in an ethical and informed manner, guided by a series of values that operate to minimize risk. Better practice guidelines supporting the development of automated systems intended to apply legislation to determine an entitlement, offer a starting point as to what considerations ought to guide deployment. These identify the substance, breadth, structural, and semantic complexity of the legislation; the remit of the authorizing agency; the transparency of decisions reached; the grounds for decision review; privacy; data accuracy; auditing; de-skilling of decision makers; and the cost of implementation and maintenance as relevant factors.[303] Building on these guidelines, we propose emerging legal AI tools are assessed with reference to the extent to which they:

- Operate to reveal rather than embed bias;
- Promote the accessibility of the legal system;
- Ensure the processes by which outcomes reached are transparent;
- Balance rather than entrench power imbalances between parties;
- Enhance rather than diminish the intelligibility of legal language;
- Sustain or improve upon the existing models of due process;
- Distribute benefits;
- Make clear the model of law imposed; and,
- Benefit from on-going monitoring and refinement.

307. AUSTRALIAN GOVERNMENT AUTOMATED ASSISTANCE IN ADMINISTRATIVE DECISION-MAKING WORKING GROUP, AUTOMATED ASSISTANCE IN ADMINISTRATIVE DECISION-MAKING BETTER PRACTICE GUIDE (2007); ADMINISTRATIVE REVIEW COUNCIL, *supra* note 164.

Furthermore, we suggest the need to[304]:

- Make the system/algorithm publicly available;
- Implement independent auditing mechanisms;
- Publish the data used to train the system;
- Categorize systems in line with an established algorithmic risk scale;
- Inform users where AI has been used in a process or decision and make clear the inputs that featured in the decision;
- Make clear who is accountable for any action that arises as a by-product of the use of the tool; and,
- Ensure that mechanisms are in place to compensate those negatively impacted by mistaken decisions.

These suggestions represent the start of a framework for evaluating AI in law, though the feasibility, practicality, efficacy, and impact of the suggestions provided will remain a work in progress as technological and regulatory infrastructures continue to evolve. Adequately capturing the full range of philosophical, practical, jurisprudential, democratic, societal, and rule of law implications that may arise as a result of AI adoption in law, and regulating these risks via appropriate governance structures, remains a critical focus for future research.

V. Conclusion

Whilst early AI efforts in the 1980s focused on the development of expert logic systems in law, as the processing power and storage capability of computers have grown, so too has the use of data-driven AI, represented by quantitative and "machine learning" methods. Studies demonstrate that it is possible to apply ML to a range of legal tasks, many of which are documented above. In adding to this literature, this paper has considered what potential AI holds as a means by which to improve the functioning of legal systems, democratize access to justice and legal services, address latent legal need, protect citizens from abuses of power, address systemic bias, and promote greater institutional accountability and transparency. Such questions are not new, and the implications they pose are the same issues with which those who create and study the law have always struggled. Nevertheless, as technology infrastructures continue to evolve,

---

304. A number of these suggestions were adapted from the work of Copeland. *See* Eddie Copeland, *10 Principles for Public Sector use of Algorithmic Decision Making*, NESTA Blog (Feb. 20, 2018), https://www.nesta.org.uk/blog/code-of-standards-public-sector-use-algorithmic-decision-making [https://perma.cc/86TM-T5F9].

increasing in complexity and sophistication and reaching further into our lives than ever before, these questions demand renewed attention.

This Paper has engaged with many of these questions and the constitutional and rule of law challenges posed by AI. We have raised concerns regarding the way in which AI systems impose a particular theoretical paradigm, limiting space for legal dialectical pluralism. We have noted a number of challenges relating to the capacity of AI to subvert the intent and purpose of the law or to fail to reflect the complexity of the expression of law. We have also considered the potential implications of AI in the context of the rule of law, observing the need for AI systems to balance a range of competing priorities and questioning their ability to do so. Finally, we have considered the pragmatic issues that are likely to shape the future of AI development in law, assessed the implications that these issues pose by reference to the likely beneficiaries of AI technologies, and offered a range of criteria that may form the basis of a more well-developed suite of safeguards that scaffold development in the field.

We conclude by emphasizing that systems capable of creating, interpreting, applying, and enforcing the law requires a transformation of what we understand as "law." The vision of law reflected in AI systems is shaped by the limits of the technology used to create that vision. Where law (whether represented by interpretation, bargaining, or enforcement) is transposed into rule-based and data-driven AI systems, nuance is supplanted by simplicity and relational representation in a manner that can operate to undermine the ends to which law is directed. Any technological representation of law and the legal obligations it spawns is necessarily bound—and in some instances irrevocably limited—by what the contemporary technology makes possible. The view of law through the eyes of AI would have been more different in the 1980s than today, not because the law has fundamentally changed, but because our ability to express law and legal relationships via technology has changed.

Law is not just process, rules, sanctions, norms, or behavior. It ought not reside in the hands of those with the resources or power to direct a self-serving agenda, but in balancing the interests of many voices in society. The consequences that arise from any transformation, whether in the form of: value-added, plurality lost, efficiency gained, flexibility preserved, costs reduced, profit made, or debate and freedom safeguarded, are relevant only in so far as they impact upon the capacity of the law to achieve (and to be seen to achieve) justice. Any implementation of AI must be able to justify a contribution to that end.

For this reason, legal AI must be evaluated in advance of deployment with reference to a framework that adequately safeguards those principles enshrined in the constitution and the rule of law. We have outlined some potential components of such a framework in Part IV, though we recognize that not all objectives are equal, or for that matter, equally achievable. It may be easier to determine the key beneficiaries of a particular technology than to project a wide range of risks into the future. Whilst we can never know with certainty what threats might accompany a particular form of legal AI in advance of deployment, we see the development of a "rule of legal AI" as an important governance mechanism guiding future development in the field and suggest on-going research in order to better illuminate its contours.