How Artificial Intelligence Will Affect the Practice of Law

Benjamin Alarie
Anthony Niblett
Albert H. Yoon[1]

Abstract

Artificial intelligence is exerting an influence on all professions and industries. We have autonomous vehicles, instantaneous translation among the world's leading languages, and search engines that rapidly locate information anywhere on the web in a way that is tailored to a user's interests and past search history. Law is not immune from disruption by new technology. Software tools are beginning to affect various aspects of lawyers' work, including those tasks that historically relied upon expert human judgment, such as predicting court outcomes. These new software tools present new challenges and new opportunities. In the short run, we can expect greater legal transparency, more efficient dispute resolution, improved access to justice, and new challenges to the traditional organization of private law firms delivering legal services on a billable hour basis through a leveraged partner-associate model. With new technology, lawyers will be empowered to work more efficiently, deepen and broaden their areas of expertise, and provide more value to clients. These developments will predictably transform both how lawyers do legal work and resolve disputes on behalf of their clients. In the longer term, it is difficult to predict the impact of artificially intelligent tools will be, as lawyers incorporate them into their practice and expand their range of services on behalf of clients.

I.   Introduction

What do lawyers do? In the context of litigation disputes, we typically think of lawyers as highly-trained and highly-skilled professionals who identify the legal issues, gather the relevant facts, and determine the likely outcome were a court to adjudicate the dispute. Lawyers exercise judgment, using their experience and intuition, to assess the merits of a case to determine the best way to proceed. Such tasks have long been viewed as ones that only highly-skilled professionals can perform. Given the importance of expert human judgment, one might suspect that the costs of such services would only increase over time.[2]

But how much of what lawyers do can be automated? Recent developments in artificial intelligence – in particular, natural language processing and machine learning – have challenged traditional conceptions of human expertise. Machines now perform increasingly complex tasks much better than humans.[3] Until recently, these tasks were held out as being the exclusive domain of humans.[4] Such tasks, it was argued, require human judgment. Driving a car, for example, was thought to be too complex for a machine. But self-driving cars are very much a reality today and experts predict they will be commonplace within the next decade. Indeed, many professions have seen the automation of some of the more mundane aspects of work. Tasks that used to require human effort – such as complex calculations, counting money, creating spreadsheets, conducting risk assessments – have all been automated in ways that not only reduce cost, but offer greater accuracy and precision. Importantly, big data has been used to aid and guide the decision-making process in industries such as banking, finance, insurance, transportation, and medicine.[5]

The practice of law is clearly not immune from these technological advances and the impact of data-driven analysis.[6] This essay discusses ways in which big data and artificial intelligence will influence the practice of law. We look to how natural language processing and machine learning can assist lawyers and perform some of the tasks currently undertaken by lawyers. John

---

[2] In this respect, legal services are analogous to health care and live music performances. *See* WILLIAM J. BAUMOL, THE COST DISEASE: WHY COMPUTERS GET CHEAPER AND HEALTH CARE DOESN'T (2013).

[3] *See, e.g.,* JERRY KAPLAN, ARTIFICIAL INTELLIGENCE: WHAT EVERYONE NEEDS TO KNOW (2016); ERIK BRYNJOLFSSON AND ANDREW MCAFEE, THE SECOND MACHINE AGE: WORK, PROGRESS, AND PROSPERITY IN A TIME OF BRILLLIANT TECHNOLOGIES (2014); MARTIN FORD, THE RISE OF THE ROBOTS (2014); Benjamin Alarie, Anthony Niblett, and Albert, *Law in the Future,* 66 U.T.L.J. 423 (2016).

[4] For an example of this line of argument in the context of legal reasoning, see Cass Sunstein, *Legal Reasoning and Artificial Intelligence: How Computers "Think" Like Lawyers*, 8 U. CHI. L. SCH. ROUNDTABLE 1, 19 (2001)

[5] *See* VICTOR MAYER-SCHOENBERGER AND KENNETH CUKIER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK 6-12 (2013); JAMES MANYIKA, ET. AL., DISRUPTIVE TECHNOLOGIES: ADVANCES THAT WILL TRANSFORM LIFE, BUSINESS, AND THE GLOBAL ECONOMY (2013); Roger Parloff, *Why Deep Learning is Already Changing Your Life*, Fortune, Sept. 28, 2016, http://fortune.com/ai- artificial-intelligence-deep-machine-learning/. Long before big data became a trend, researchers – particularly in the field of medicine – had observed the superiority of algorithmic to human judgment in a number of areas. *See* William M. Grove and Paul E. Meehl, *Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The Clinical-Statistical Controversy,* (1996) 2 PSYCH. PUBLIC POL'Y & LAW 293**.**

[6] For a more general treatment of how the legal services is being disrupted by technology, *see* RICHARD SUSSKIND, TOMORROW'S LAWYERS: AN INTRODUCTION TO YOUR FUTURE (2014) and RICHARD SUSSKIND AND DANIEL SUSSKIND, THE FUTURE OF THE PROFESSIONS: HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS (2016); and Raymond H. Brescia, et al., *Embracing Disruption: How Technological Change in the Delivery of Legal Services Can Improve Access to Justice*, 78 ALBANY L. REV. 553 (2015)

McGinnis and Russell Pearce have argued that machine intelligence will cause a "great disruption" in the market for legal services.[7] They posit five areas of legal services will be affected: discovery, legal search, document generation, brief generation, and prediction of case outcomes.[8] In this essay, we touch on each, but focus primarily on prediction. In particular, we focus on the ability of machines to assess the merits of a legal case.

Litigation is fundamentally about "bargaining in the shadow of the law",[9] where lawyers resolve disputes by persuasively positing what would happen if a court were to consider and decide the matter. These prognostications, however carefully reasoned, are often biased, both by the clients' desire to win and by their lawyers' own experiences. Machine learning has the potential to provide more objective predictions of how courts will decide discrete legal issues.[10] Once given facts relevant to the question, a machine can situate these facts within the domain of applicable legal precedents. In addition to being less susceptible to various kinds of biases, machines do not suffer from other problems affecting human lawyers exercising such judgment. Algorithms can generate predictions that others can replicate. Moreover, algorithms do not tire. Computers do not need to take time off.

Such technological advances make it possible for parties to litigate more effectively, before, during, and after trial, by predicting legal outcomes with greater accuracy and consistency, and at lower cost. Our discussion is framed with reference to broader changes in the legal services market and with reference to access to justice. In the aftermath of the 2008 recession, lawyers across diverse practice settings – large firms, small firms, government, public interest, corporations – have explored ways to provide more cost-effective legal representation. Lawyers in private practice environments find that clients are less willing to accept the indeterminacy of hourly billing. Government and public-interest organizations, in response to budgetary austerity, are compelled to explore ways to do more with less. The incentive to adopt cost-saving and precision-increasing technologies is and will continue to be great.

This essay proceeds as follows: Part II provides a brief overview of the legal profession. Part III describes the evolution of information in legal practice. Part IV introduces and briefly describes recent technological advances in artificial intelligence, specifically natural language processing and machine learning. Part V provides an overview of how these advances can be used to automate parts of a lawyer's job. Part VI discusses how these tools will impact upon the provision of legal services, specifically relating to efforts to manage litigation risk. Part VII takes a longer-term view, asking whether the adoption of artificial intelligence in law will be evolutionary or revolutionary. Part VIII concludes.

---

[7] John O. McGinnis and Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services,* (2014) 82 FORDHAM L. REV. 3041.
[8] *Id.*
[9] *See* Robert N. Mnookin and Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case for Divorce,* 88 YALE L.J. 950 (1979).
[10] *See, e.g.*, Daniel Martin Katz, *Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 936 (2013); Harry Surden, *Machine Learning and Law*, 89 UNIV. OF WASH. L. REV. 87, 102-110 (2014); Daniel Martin Katz, Michael Bommarito, and Josh Blackman, *A General Approach for Predicting the Behavior of the Supreme Court of the United States,* PLoS ONE 12(4): e0174698 (2017).

II.   The evolution of the legal profession

The legal profession is a complex labour market. It consists of several moving parts: solo or small firm practitioners, large law firms, government, public interest firms, judges, among others. Our focus in this section is on the private provision of legal services by law firms, which accounts for the majority of the legal services provided in North America. Law firms have traditionally tended to provide legal services on a billable hour basis and, at large law firms, to leverage large numbers of associates for each law firm partner.

The private practice of law, at least in Canada and the United States, has three distinctive features that have supported the perpetuation of the billable hour. The first is that legal practice has tended to increase in complexity over time. Courts steadily render new decisions, adding to the corpus of common law. At the same time, legislatures continue to pass new legislation, more often to modify or amend rather than supersede existing legislation. Because lawyers must keep abreast of both judicial and legislative developments, it follows that lawyers will tend to increasingly narrow their scope of expertise, particularly as the returns to specialization increase.[11]

The second feature is the unpredictability of the time required to complete a legal task. It depends both on the question of law and the relevant facts, which vary from one case to the next. Situations that give rise for legal representation are rarely answered definitively by existing precedent or statutes. Lawyers must extrapolate from relevant cases, drawing analogies and identifying distinctions to best promote their clients' interests. Moreover, completion of one legal task often gives rise to imperfectly anticipated subsequent legal tasks as the position and interests of clients evolve over time.

Third, in many areas of the law, the nature of legal practice is client-oriented. A client's relationship may nominally be to the law firm as a whole, but it is effectively tied to individual lawyers within the firm. Because of an individual lawyer's familiarity with the specific affairs of particular clients, an individual lawyer cannot freely hand off many assignments to other lawyers within the firm. Moreover, she may have a disincentive to do so, risking that the client prefers work done by other lawyers within the firm.

In the aftermath of the 2008 global recession, clients of large law firms began to challenge the traditional model of hourly billing. The economic downturn gave clients of law firms leverage to demand more cost-effective representation.[12] Law firms began experimenting with alternative approaches, such as flat rates and contingency fees, in new contexts. These new approaches were typically less expensive for clients than hourly billing.[13]

---

[11] *See* Luis Garciano and Thomas N. Hubbard, *Specialization, Firms and Markets: The Division of Labour Between and Within Law Firms,* 25 J. L. ECON. ORG. 339 (2009).
[12] *See* Catherine Ho, "Law Firms Look For Alternatives to the Billable Hour?" WASH. POST, April 15, 2012 (available at https://www.washingtonpost.com/business/economy/law-firms-look-for-alternatives-to-the-billable-hour/2012/04/15/gIQAeyW9JT_story.html?utm_term=.89ebd0767473).
[13] *See* Catherine Ho, "Is This the Death of Hourly Rates at Law Firms?" WASH. POST, April 23, 2014 (available at https://www.washingtonpost.com/business/capitalbusiness/is-this-the-death-of-hourly-rates-at-law-firms/2014/04/11/a5697018-be97-11e3-b195-dd0c1174052c_story.html?utm_term=.9ef2d7b92951).

Client demands also compelled law firms to rethink their institutional economic model. Law firms historically relied on a two-tiered model of partners and associates. Partners had ownership stakes in the firm; associates did not. In an effort to reduce operating expenses, law firms moved away from this traditional model.[14] The emergence of non-equity partners (an intermediate step for associates in their initial promotion to partner) and increased numbers of *of-counsel* positions has transformed the "partnership tournament" into an "elastic tournament" with several gradations, with fewer paths to equity partnership. The modern law firm now has fewer equity partners at the top, and several rungs below to reflect the increasing variety of lawyer roles in the firm, ranging from non-equity partners, of-counsel, non-partner-track associates,[15] contract lawyers,[16] and outsourced lawyers.[17]

Client demand for lower legal expenses is only part of the story behind the evolution of private legal practice. It provided the necessary impetus for law firms to transform their business model. Technological advances provide the tools for law firms to facilitate this transition, and make it not only financially viable but also potentially lucrative for a firm's equity partners. These recent advances, while dramatic in their effect, are merely the latest transition in the technical progression of legal information.

There are, of course, other forces in the legal service markets that may hinder the adoption of new technologies. In particular, legal scholars have suggested that because lawyers self-regulate their own market, they will be able offer strong resistance to the development, use, and integration of machines into legal practice.[18] McGinnis and Pearce, on the other hand, contend that these regulatory barriers will not prevent machine intelligence from "undermining lawyers' monopoly".[19] They argue that the unauthorized practice laws are rarely applied successfully to machine intelligence products.[20] Nonetheless, fears of the impact upon machine intelligence may impinge upon the lawyers' monopoly may provide impetus for rent seeking opportunities by lawyers who respond by strengthening these laws.

In the next section, we discuss how the ability to access to legal information has evolved over time.

---

[14] *See, e.g.,* Peter Lattman, "Mass Layoffs at a Top-Flight Law Firm," N.Y. TIMES, June 24, 2013 (available at https://dealbook.nytimes.com/2013/06/24/big-law-firm-to-cut-lawyers-and-some-partner-pay/?_r=0); Dan Slater, "At Law Firms, Reconsidering the Model for Associates Pay," N.Y. TIMES, April 1, 2010 (available at https://dealbook.nytimes.com/2010/04/01/at-law-firms-reconsidering-the-model-for-associates-pay/).

[15] *See* Catherine Rampell, "At Well-Paying Law Firms, a Low-Paid Corner," N.Y. TIMES, May 23, 2011 (available at http://www.nytimes.com/2011/05/24/business/24lawyers.html).

[16] *See* Vanessa O'Connell, "Lawyers Settle . . . for Temp Jobs," Wall St. J., June 15, 2011 (available at https://www.wsj.com/articles/SB10001424052702303714704576383641752966666).

[17] *See* Heather Timmons, "Legal Outsourcing Firms Creating Jobs for American Lawyers," N.Y. TIMES, June 2, 2011 (available at http://www.nytimes.com/2011/06/03/business/03reverse.html).

[18] *See, e.g.,* Larry E. Ribstein, *The Death of Big Law*, 2010 WIS. L. REV. 749, 807; Ray Worthy Campbell, *Rethinking Regulation and Innovation in the U.S. Legal Services Market*, 9 N.Y.U. J.L. & BUS. 1 (2012); Gillian Hadfield, *Legal Barriers to Innovation: The Growing Economic Cost of Professional Control Over Corporate Legal Markets*, 60 STAN. L. REV. 1689, 1724-5 (2008).

[19] McGinnis and Pearce, *supra* note 5, at 3057-64. See also BENJAMIN H. BARTON, A GLASS HALF FULL LOOK AT THE CHANGES IN THE AMERICAN LEGAL MARKET (2015).

[20] McGinnis and Pearce, *supra* note 7, at 3061-4.

III. The evolution of legal information

The provision of legal services has evolved, in part, due to changes in access to legal information. In this section, we briefly discuss three epochs of access to legal information. The first stage is the use of *analog* materials. These include case reporters, law journals, casebooks, textbooks, loose-leaf services, and legislation in hard copy. The second stage of the evolution involves the use of *digital* materials and libraries, gaining traction in the 1990s with the development of the CD-ROM. Nowadays, law firms have almost all made the transition to digital legal resources. The third stage of legal information will be the focus of the remainder of this essay. This stage involves the use of predictive tools to sort through noise and more easily locate and synthesize relevant information. This information includes objective determinations of the likely outcomes of litigation. We refer to this third stage of legal information and legal research as *computational*.

***Analog*:** Consider how laws were maintained, learned, and researched prior to the digitization of legal materials. In ancient times, laws that were codified in writing and promulgated, such as those in the times of Hammurabi and Justinian I, were rare. In ancient Rome and Greece, lawyers primarily represented their clients orally in appearances before the courts.[21] Memorialization of the law into written word, when it occurred, was costly, given that it had to be reproduced by hand. Little changed throughout the centuries. Indeed, the above situation could easily describe the way in which law was practiced in England in the 16th and 17th centuries. The advent of the printing press made it possible to disseminate this information more broadly.[22] But this technology took centuries to impact how law was produced and consumed. Judicial decisions were placed in published volumes, as were legislative acts, regulations, and academic and practice-based commentaries.

The availability of hard copy materials defined the universe of relevant materials for all stakeholders: judges, lawyers, and litigants. As Oliver Goodenough has noted, Abraham Lincoln's practice of law was constrained by the number of books he could carry on his horse.[23] A central limitation was not just the physical constraints of how much one could carry, but rather that dissemination of these materials remained limited. These materials were typically housed in law libraries, which were reserved for law schools, law societies, and the wealthiest law firms. Many lawyers had only limited access to these references, which in turn placed them – and their clients – at a competitive disadvantage. Of course, even for lawyers with access, the quality of representation was only as good as their ability to absorb and effectively synthesize these materials. Access was necessary, but by no means sufficient, for quality legal representation.

---

[21] For an interesting discussion of the legal system in ancient Greece, see ADRIAAN LANNI, LAW AND JUSTICE IN THE COURTS OF CLASSICAL ATHENS (2008).

[22] *See generally,* ELIZABETH L. EISENSTEIN, THE PRINTING PRESS AS AN AGENT OF CHANGE (1982).

[23] Keynote speech at Codex FutureLaw 2015 (at http://conferences.law.stanford.edu/futurelaw2015/).

Digital: In 1976, Westlaw and Lexis began offering computer-assisted legal research.[24] Instead of relying on physical books for their reference needs, lawyers could search through these materials first through the use of a terminal and, later on, a personal computer. The texts of these references were in electronic format, meaning that lawyers could use keyword searches to find relevant sources. Keyword searches enabled lawyers to more quickly survey vast libraries for relevant materials, something that would take much longer using the traditional, manual approach of looking through reference indices and then at each source one by one..

As with its analog predecessor, digital referencing suffered from the problem of limited availability. Access to digital libraries can be costly, and remained out of reach for many lawyers, particularly sole practitioners. Importantly, as with analog provision of legal references, digital access enabled lawyers to research faster, but still did not guarantee quality legal representation. A keyword search poses the risk of being under- and over-inclusive in defining the scope of relevant materials. In the former situation, a search may result in the exclusion of relevant materials, simply because those materials do not contain the keywords in the search. In the latter situation, a search may include irrelevant materials, because the reference includes the keyword but is otherwise not germane. As a result, digitized legal research could be incomplete, costly, and in many instances, both.

Lawyers today conduct most of their legal research online. Those with greater resources use proprietary data providers such as Westlaw, Lexis, or Bloomberg; others rely on publicly available resources such as Justia (United States) or CanLII (Canada); or general search engines such as Google.[25] Most recent government documents – e.g., judicial opinions, legislation, regulations – are available online, although proprietary databases have more powerful searching capabilities because providers have invested in structuring aspects of the data.[26]

Computational: We envision the emergence of *computational* legal research tools based on artificial intelligence. Such tools can be used to access relevant information with greater speed, lower cost, and higher accuracy. The predictive power of these tools will transform the way lawyers approach research and litigation.[27] In 1897, Oliver Wendell Holmes, Jr., wrote, "For the rational study of the law, the black-letter man may be the man of the present, but the man of the future is the man of statistics and the master of economics."[28] The intuition underlying his application is that of prediction: "prophecies of what the courts will do in fact . . . are what I mean by the law."[29] The development of artificially intelligent tools will facilitate such prediction.

---

[24] *See* James A. Sprowl, *Computer-Assisted Legal Research: Westlaw and Lexis,* 62. AMER. BAR ASSOC. J. 320 (1976).
[25] In addition, many courts now make their published decisions available online.
[26] For example, the proprietary data providers routinely tag legal documents, such that it is possible for users to narrow their search along specific parameters (e.g. title, citation, court level) or create structured parameters (e.g., synopses).
[27] *See*, McGinnis and Pearce, *supra* note 7; Katz, *supra* note 10; Surden, *supra* note 10.
[28] Oliver W. Holmes, *The Path of the Law,* 10 HARV. L. REV. 457, 469 (1897). For a nuanced interpretation of the implication of Holmes's statement, see Hildebrandt, Mireille, Law As Computation in the Era of Artificial Legal Intelligence. Speaking Law to the Power of Statistics (June 7, 2017). Available at SSRN: https://ssrn.com/abstract=2983045.
[29] *Id.*

In the next section, we provide a (very brief) overview of some of the basic concepts in artificial intelligence. In the section that follows, we provide an overview of how these tools can and will be applied in legal practice.

IV. The emergence of artificial intelligence

In this section, we briefly introduce specific concepts in artificial intelligence that will likely spur the development of computational legal tools. We focus on natural language processing and machine learning. Proponents of artificial intelligence contend these developments offer the promise of technology to improve how humans engage with the world. For its detractors, artificial intelligence represent technical threats whereby machines will colonize tasks that were historically the domain of humans.[30]

"Artificial intelligence" is a notoriously difficult term to define.[31] There is little agreement as to what the term means. The technological advances that fall under the rubric of artificial intelligence depend on the point in time that one seeks to define its meaning. The concept changes over time. For the purposes of this essay, we will address artificial intelligence as a somewhat nebulous branch within computer science that seeks to build machines capable of what humans would regard as "intelligent" behaviour. Of course, what is meant by intelligent behaviour is also difficult to pin down.

Natural language processing and machine learning are specific subfields or applications of artificial intelligence. Natural language processing examines the use words and phrases to draw connections within and across written or spoken language. Machine learning is the process of enabling computers to learn to optimize certain tasks without the benefit of explicit rules-based programming. Successful applications of machine learning include chess and, more recently, Go.

Natural language processing enables machines to adapt when evaluating text. A keyword search is literal in its approach, looking for exact words or phrases. By contrast, natural language processing allows the user to identify materials that are likely relevant to her search, even if the materials do not contain words or phrases expressly within her list of keywords. Natural language processing applies to both retrieving information (e.g., identifying the relevance of a document) or extracting information (determining the key terms from a document). Given its advantages, natural language processing tools have largely replaced keyword searches in many fields. Search engines, speech-to-speech translation, and intelligent assistants such as Siri are built with natural language processing technology for the user's benefit.[32]

---

[30] For example, Stephen Hawking recently commented, "I believe there is no deep difference between what can be achieved by a biological brain and what can be achieved by a computer." *See* Rory Cellan-Jones, "Stephen Hawking – Will AI Kill or Save Humankind?", BBC News, Oct. 20, 2016 (available at http://www.bbc.com/news/technology-37713629).

[31] For an interesting discussion of the definition and history of the term "artificial intelligence", *see* Kaplan, *supra* note 3, at 1-48.

[32] For a discussion of the recent developments of natural language processing, see Julia Hirschberg and Christopher D. Manning, *Advances in Natural Language Processing,* 349 SCIENCE 261 (2015).

Machine learning represents a dramatic advance in the evolution of artificial intelligence.[33] In tackling earlier challenges – e.g., defeating a grandmaster in chess[34] – artificial intelligence used computational brute force and position evaluation functions to plot out the likely course of the game and to optimize for the selection of the next move. As artificial intelligence gravitated to more complex problems involving an undefined or infinite outcome space, machine learning experts developed tools that identify nuanced patterns in data, beyond what highly skilled, experienced workers could reasonably construct on their own. Recently, systems built using machine learning have convincingly defeated Go masters[35] and World Series of Poker champions.[36]

Both natural language processing and machine learning are seemingly capable of processing large amounts of seemingly unstructured data. The actual process is more nuanced for each. Natural language processing looks for structure whenever available. For example, it tags words for its parts of speech (noun, verb, adjective, etc.); or draws connections between references, even if the references use different terminology. Machine learning synthesizes large amounts of data – often unstructured – by identifying the components that it observes, and developing algorithms that maximize its predictive accuracy. Machine learning's agnostic approach – choosing an algorithm that maximizes predictive accuracy independent of underlying theory – enables it to leverage connections between and among references, even those that are implied rather than express.

V. Computational legal information

In this section, we consider how lawyers may adopt tools based on natural language processing and machine learning. Consider the tools that lawyers already use to cut down on much of the mundane work of junior associates. One such example is e-discovery. In the days of traditional discovery, the process was highly labour-intensive, requiring a small army of lawyers to manually read through boxes of documents in search of information relevant to litigation. Because most legal documents now originate in electronic format, the discovery process can now be conducted with a fraction of the time, expense, and need for lawyers, with even greater accuracy.[37] Automated basic keyword searches alone offer vast improvement over traditional means of discovery of lawyers manually reading over documents; the use of natural language processing to assess relevance provides an additional improvement. As a result, the discovery

---

[33] *See, e.g.,* Kaplan, *supra* note 3, at 27-39.

[34] For a description of IBM's efforts, see FENG-HSIUNG HSU, BEHIND DEEP BLUE: BUILDING THE COMPUTER THAT DEFEATED THE WORLD CHESS CHAMPION (2004).

[35] *See* Christof Koch, "How the Computer Beat the Go Master," SCIENTIFIC AMERICAN (March 19, 2016) (available at https://www.scientificamerican.com/article/how-the-computer-beat-the-go-master/); *see also* Paul Mozor, "Google's AlphaGo Defeats Chinese Go Master in Win for A.I.," N.Y. Times, May 23, 2017, at B3 (available at https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html).

[36] *See* Tonya Riley, "Artificial Intelligence Goes Deep to Beat Humans at Poker," SCIENCE (March 3, 2017) (available at http://www.sciencemag.org/news/2017/03/artificial-intelligence-goes-deep-beat-humans-poker).

[37] *See* John Markoff, "Armies of Expensive Lawyers, Replaced by Cheaper Software," N.Y. TIMES, March 4, 2011, at A1 (available at http://www.nytimes.com/2011/03/05/science/05legal.html) (discussing how discovery for a case involving 6 million documents cost $2.2 million in 1978, compared with a 2011 case involving 1.5 million documents for a cost of less than $100,000.

process is now largely outsourced to e-discovery firms, and lawyers now spend less than 5 percent of their time on basic document review.[38]

Natural language processing has also changed how lawyers approach forms and contracts. A plethora of start-up companies work on contract review and analysis. The companies market software that helps parties develop contractual language that avoids many of the pitfalls of drafting, such as ambiguous conditions and omission of key provisions or terms.[39] The software also facilitates the interpretation of existing contracts, in response to or in anticipation of litigation. Other companies provide analytic software tools to give structure to unstructured data (e.g., email) relevant to existing or prospective litigation.[40] In the area of bankruptcy, a firm uses natural language processing to identify existing published decisions most relevant to a client's case, and then follows up with a written memo the next business day.[41]

The focus of this section, however, is prediction. In particular, we focus on how predictive tools can be used to assess the merits of a case. For lawyers, evaluating how a litigant's current case maps onto the existing set of published court decisions can present several challenges. A discrete legal issue can be shaped by hundreds or thousands of judicial decisions. Also, each dispute is unique, often in ways imperceptible to even the most experienced lawyers. Lawyers, however well trained, are limited in their ability to evaluate all relevant decisions and are prone to personal biases. Machine learning tools can help overcome these biases, by providing a more objective and independent prediction of the likely outcome of a particular litigation strategy.

A central challenge for using artificial intelligence in law is the unstructured nature of legal data. Take, for example, judicial opinions. While opinions follow a general form – recitation of facts, discussion of relevant case law, and application of law to the facts – judges are highly individualistic in the way they present this information. They vary considerably with one another in their writing style, including grammar and diction.[42] When dissenting with one another, they use different legal precedents to support their analyses.[43]

Judges' personal approaches to opinion writing can present obstacles to tools that seek to extract relevant information. Search terms can be noisy, generating under- and over-inclusive results. The highly contextualized writing can make it difficult to identify patterns within and across opinions. Accordingly, algorithms relying on natural language process may yield imprecise results, and predictably more so as the complexity of the legal question increases.

It is, of course, possible to transform such unstructured data into structured data. This process involves identifying a discrete question, the corpus of relevant documents, and the most relevant

---

[38] *See* Dana Remus and Frank Levy, *Can Robots be Lawyers? Computers, Lawyers and the Practice of Law, forthcoming*, GEO. J. LEG. ETHICS (2017). *See also*, McGinnis and Pearce, *supra* note 7, at 3047-8.
[39] *See, e.g.,* Kira Systems (kirasystems.com).
[40] *See, e.g.,* Brainspace (brainspace.com).
[41] *See* Steve Lohr, "A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet." N.Y. TIMES, March 19, 2017 (available at https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html?_r=0).
[42] *See* Jeffrey S. Rosenthal and Albert Yoon, *Detecting Multiple Authorship of U.S. Supreme Court Legal Decisions Using Function Words*, 5 ANNALS OF APPLIED STATISTICS 283 (2011).
[43] *See* Anthony Niblett and Albert H. Yoon, *Friendly Precedents, 56* WM & MARY L. REV. 1789 (2016); Anthony Niblett and Albert H. Yoon, *Judicial Disharmony: A Study of Dissent,* 42 INT'L REV. L. & ECON. 60 (2015).

features of the information necessary to extract from the documents. This process may be labour intensive, requiring the inputs of lawyers, machine learning experts, and developers to complete each stage. This process, while costly, is front-loaded. The cost of applying machine learning to this data – now structured – is modest by comparison. And once completed, this digitization of law allows for unlimited economies of scale, available to any individual without reducing its availability to anyone else.[44]

Blue J Legal, a Toronto-based startup,[45] applies machine learning to predict how courts would decide legal questions. The start-up began with tax law: a complex and dense area of law often based on vague standards. Blue J Legal uses machine learning to build tax law *classifiers*, based on published Supreme Court of Canada, Federal Court of Appeal, and Tax Court of Canada decisions, in which the court makes a binary determination for a specific legal question. Each classifier is structured to answer how would a court likely decide, based on a given set of facts.

The first step is to identify a fact-intensive question of law. This question generates a defined universe of cases, and factors most relevant to the court when deciding the question. One example is a worker classifier that determines whether a worker is an employee or independent contractor for tax purposes.[46] In each case, the court considers a number of different factors related to the degree of control the employer exercises over the worker, the ownership of tools or equipment, the financial risk incurred by the worker, and the integration of the worker into the hirer's business.[47] Constructing the relevant questions is a time-intensive, iterative process.

The next step is to code every published decision in accordance with the aforementioned factors. This process turns unstructured data (the text of judicial opinions) into structured data (discrete information from the aforementioned questions in the form of variable data). Blue J Legal uses machine learning technology to generate a predictive algorithm. These tools identify connections among the variables that are difficult to specify using more traditional statistical techniques. After the user answers the relevant questions, the classifier provides a written report containing the following: 1) the likely outcome; 2) the percentage likelihood of the outcome; 3) a detailed memo explaining the rationale for the outcome, based on the specific answers; 4) the leading court decisions for the classifier; and 5) the precedents that are most similar, based on the user's responses to the questions.

Of course, such tools will not be available for every question within specific sub-fields of law. The temptation of artificial intelligence is to view it as a proverbial hammer where all the legal questions are nails. The law does not fit this paradigm. It is inappropriate in instances where the court's determination of a legal question does not lend itself to an identifiable set of factors, or where insufficient data exist. Future developments may well be able to surmount these current challenges of limited data and inchoate context; time will tell.

---

[44] Such economies of scale are typical of disruption by software. *See, e.g.,* Marc Andreessen, "Why Software Is Eating the World?", *The Wall Street Journal,* Aug 20, 2011.
[45] In the interests of full disclosure, we reiterate that the three co-authors are also co-founders of Blue J Legal.
[46] The early progress of the Blue J Legal project is documented in Benjamin Alarie, Anthony Niblett, & Albert H. Yoon, "Using Machine Learning to Predict Outcomes in Tax Law", (2016) 58 *Canadian Business Law Journal* 231.
[47] Under Canadian tax law, these factors are based on the "total relationship" test as set out in *Wiebe Door Services Ltd. v. M.N.R.*, [1986] 3 F.C. 553 (FCA).

VI. The impact of computational legal information in the short term

What is the role of the lawyer in a world with computational legal tools? Because artificially intelligent tools can cut through the morass of unstructured legal information and situate new cases in the context of relevant precedent, it is tempting to raise the possibility that litigants could simply use this technology directly and bypass a lawyer altogether. This approach, however, misconstrues the purpose of both this technology and the role of the lawyer. These technological tools are merely the next generation of legal information. They do not purport to offer legal advice in the same way that lawyers do. Accordingly, these tools do not supplant what lawyers fundamentally do on behalf of their clients.

A lawyer does not merely provide answers to legal questions. She identifies the appropriate question (or likely, multiple questions) to answer. Moreover, her task is to advise a course of action that goes beyond the specific and immediate legal question, to consider the client's overall circumstances and interests. A lawyer who uses artificial intelligence will still need to exercise her reasoned judgment to evaluate the accuracy of the purported facts, as well as the appropriateness of the technology in resolving the legal dispute. This reasoned judgment is something that artificial intelligence – whether through natural language processing or machine learning– is not, as yet, able to perform.

Medicine provides an instructive analogy. Individuals can avail themselves of an abundance of information online: e.g., ailments, diseases, treatments. In addition, apps and websites enable individuals to self-diagnose.[48] While this technology can help inform individuals, it cannot replace a physician, who possesses a comprehensive understanding of medicine as well as of the patient (often beyond what the patient herself is aware of). Accordingly, while individuals may achieve better health outcomes using self-help measures than foregoing all efforts altogether, they benefit most by seeing a physician who uses artificial intelligence tools to facilitate medical diagnoses.

If legal classifiers are able to provide accurate predictions for discrete questions of law, lawyers can take advantage of them in a variety of ways. They can provide plaintiffs and defendants with timely, objective assessments of their claim. More importantly, each side can benefit from the same methods of generating these assessments. This transparency in the law will help parties at every stage of litigation, from early settlement talks to trial preparation to appeals. These classifiers can help parties focus on pivotal fact-based areas of disagreement, which can be more easily overcome with evidence.

Government can also use legal classifiers to help evaluate claims and manage litigation risk. At the early stages of a prospective issue, any government worker responsible for rendering decisions can use legal classifiers to ensure that they are providing consistent, accurate advice.[49]

---

[48] See, e.g., the Symptom Checker page on WebMD.com (available at https://symptoms.webmd.com/#introView, last checked on October 22, 2017).

[49] Consider how tax laws are regulated and enforced. To effectively enforce these laws requires an agency to possess a strong understanding of how a court would interpret them. Humans working within the agency – e.g., agents, auditors, lawyers – make these decisions, and they routinely differ from one another when deciding the same legal

For administrative claims, the government can use the classifier to determine whether they should challenge the claim or simply settle. And for cases that enter formal litigation, the government can use a classifier to develop strategy, identifying the strongest and weakest parts of their legal argument, and discern the relative importance of various factual claims. In so doing, the government can deploy limited litigation resources more efficiently.

Artificial intelligence will help lawyers in three dimensions. First, lawyers are able to do more in the same amount of time. Their increased efficiency can enable them to perform additional work on behalf of existing clients, or perhaps help more clients. Second, because artificial intelligence enables them to perform tasks in minutes that historically took hours, lawyers are less tethered to working in large law firms. Smaller firms can perform as capably as larger firms, since artificial intelligence increasingly obviates the need for massive labour power emblematic of large law firms. Third, lawyers may be able to broaden rather than narrow their areas of specialization. Unshackled from the daunting task of reading a growing corpus of materials, lawyers can use artificial intelligence to maintain areas of expertise, and develop new ones.[50]

As technology continues to develop, regulatory issues will inevitably arise. The first centers around evolving standards of professional responsibility lawyers owe their clients. Looking further ahead, as predictive tools become more accurate, lawyers may not only come to rely on artificially intelligent tools, but may be legally required to consult these tools as part of their due diligence. A lawyer, by failing to access the most accurate legal information or failing to use predictive tools to better assess how a case will develop, may expose herself professional liability in tort. A displeased client may be able to recover losses if she loses a case that was brought or contended based on the hunch of a lawyer, who either failed to consult appropriately sophisticated tools or ignored the predictive information. In this way, a lawyer may be held to be negligent in the same way that a doctor is held negligent for damage that may accrue as a result of either not ordering a blood test or for ignoring the results of a blood test.

VII. The longer term impact: Evolution or revolution?

We have, thus far, suggested that lawyers can use machine learning technology to help cut through noise and better understand the objective merits of a client's case. But, in the context of the long term, this may grossly understate the impact of artificial intelligence on both the law and the way in which legal services are provided. The longer-term developments may represent more of a revolution rather than a mere evolution in legal information. In this section, we explore two possibilities for such a revolution.

First, we have assumed that artificially intelligent tools will substitute for human tasks that lawyers currently undertake. But many revolutionary changes in technology are not simply

---

matter. The volume of relevant court decisions can be a hurdle for regulators. Even after several hours of research, the agent's decision may be incorrect or inconsistent with other agency decisions. *See* Dean Beeby, "Revenue Canada's Call Centres Giving Bad Tax Advice: Report," CBC News, February 8, 2015 (available at http://www.cbc.ca/news/politics/revenue-canada-s-call-centres-giving-bad-tax-advice-report-1.2946998).
[50] *See* Albert H. Yoon, *Cognitive Computing and the Future of the Legal Profession,* 66 *U. Toronto L.J.* 456, 471 (2016).

substitutes for existing processes. Consider how electricity changed the manufacturing process. In the short term, some factories simply replaced steam powered engines with motors powered by electricity. These manufacturers did not experience a great productivity boon. The explosion in productivity came later when firms restructured their systems, architecture, logistics, and personnel to take advantage of the new technology.[51]

A similar phenomenon occurred when computers were introduced into the workplace. Taking the old system and simply inserting computers did little to improve productivity.[52] Thus, the true benefits of artificially intelligent tools in the legal profession may be realized only once lawyers completely rethink the provision of legal services. Rather than these tools simply doing what lawyers do, merely better and faster, artificial intelligence may provide impetus for a complete overhaul in the way legal services are provided. For example, as legal work becomes less costly and more efficient, lawyers can charge less for discrete tasks while simultaneously accomplishing more work in the same amount of time. In effect, artificial intelligence may increase access to civil and criminal justice in a way that benefits lawyers and clients alike.

Second, big data and artificial intelligence will likely change the way that law is produced and consumed and, indeed, perhaps even change the nature of law itself. In other pieces, we have investigated how regulators can use algorithms to determine likely outcomes, which can foster greater certainty, transparency, and openness in regulation.[53]

The judiciary may use such predictive algorithms as well. Consider how machine learning may be used in the context of bail decisions in the United States. Judges determine bail based in large part on whether they believe the defendant will violate the terms of bail if released. Historically, judges made these decisions based on their own intuition, with little oversight or even awareness of how their fellow jurists were deciding bail. Some jurisdictions now allow judges to base their bail decisions on algorithms that predict the likely risk of a defendant skipping bail or committing a crime while out on bail. These algorithms use the universe of available bail decisions.[54] Researchers in economics and computer science have applied machine learning to this field to illustrate the benefits of machine learning algorithms over human decision-making. A recent study of bail hearings in New York found that machine learning algorithms may reduce crime by 25% without increasing the jail population or reduce the jail population by 42% without increasing the crime rate.[55]

---

[51] *See,* Paul David, *The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox,* (1990) 80 AM. ECON. REV. 355; TIM HARFORD, FIFTY INVENTIONS THAT SHAPED THE MODERN ECONOMY (2017), at 91-5.
[52] Erik Brynjolfsson and Lorin Hitt, *Beyond Computation: Information Technology, Organizational Transformation and Business Performance,* (2000) 14(4) J. OF ECON. PERSPECTIVES 23; and, more generally, ERIK BRYNJOLFSSON AND ANDREW MCAFEE, MACHINE, PLATFORM, CROWD (2017).
[53] Benjamin Alarie, Anthony Niblett, and Albert H. Yoon, *Regulation by Machine,* (2017 forthcoming), J. OF MACHINE LEARNING RESEARCH.
[54] *See* Wisconsin v. Loomis, 2016 WI 68 (2016) (upholding state judges' reference to an algorithm in determining sentencing).
[55] *See* John Kleinberg, et al, *Human Decisions and Machine Predictions*, NBER Working Paper (Feb. 2017) (available at http://www.nber.org/papers/w23180).

Other work has explored how other lawmakers can use artificial intelligence to make laws that are better tailored to individual circumstances, blurring the line between rules and standards and providing for increased personalization of law.[56] In such a world where specific laws are highly personalized and communicated to individuals directly, the role of a lawyer may transform to one of a policy advisor, rather than an advisor of individual clients.

VIII. Conclusion

This essay examines recent developments in artificial intelligence that enable lawyers to make objective, accurate predictions of discrete legal questions. These developments will transform how lawyers both do legal work and resolve disputes on behalf of their clients. In the short term, the development of such predictive tools will improve transparency and empower lawyers to work more efficiently, deepen and broaden their areas of expertise, and provide greater access to justice and more value to clients. In the longer term, it is unclear exactly how dramatic the impact of artificially intelligent tools will be. Such predictions are difficult for us to make. We are, after all, only human.

---

[56] See, e.g., Anthony J. Casey & Anthony Niblett, *Self-Driving Laws,* (2016) 66 U.T.L.J 429; Anthony J. Casey & Anthony Niblett, *The Death of Rules & Standards*, (2017) 92 IND. L. REV. 1401; Omri Ben-Shahar & Ariel Porat, *Personalizing Negligence Law*, (2016) N.Y.U. L. REV. See also, John O. McGinnis and Steven Wasick, *Law's Algorithm,* 66 FLA. L. REV. 991 (2015).