NOT FOR QUOTATION
COMMENTS APPRECIATED

CAN ROBOTS BE LAWYERS?
COMPUTERS, LAWYERS, AND THE PRACTICE OF LAW

Dana Remus and Frank Levy†

We assess frequently-advanced arguments that automation will soon replace much of the work currently performed by lawyers. Our assessment addresses three core weaknesses in the existing literature: (i) a failure to engage with technical details to appreciate the capacities and limits of existing and emerging software; (ii) an absence of data on how lawyers divide their time among various tasks, only some of which can be automated; and (iii) inadequate consideration of whether algorithmic performance of a task conforms to the values, ideals and challenges of the legal profession.

Combining a detailed technical analysis with a unique data set on time allocation in large law firms, we estimate that automation has an impact on the demand for lawyers' time that while measureable, is far less significant than popular accounts suggest. We then argue that the existing literature's narrow focus on employment effects should be broadened to include the many ways in which computers are changing (as opposed to replacing) the work of lawyers. We show that the relevant evaluative and normative inquiries must begin with the ways in which computers perform various lawyering tasks differently than humans. These differences inform the desirability of automating various aspects of legal practice, while also shedding light on the core values of legal professionalism.

INTRODUCTION

On March 14, 2011, a *New York Times* headline read: "Armies of Expensive Lawyers, Replaced by Cheaper Software."[1] In the article, *Times* technology reporter John Markoff

† Professor of Law, University of North Carolina School of Law; and Professor Emeritus, Department of Urban Studies and Planning, Massachusetts Institute of Technology and Research Associate, Department of Health Care Policy, Harvard Medical School. The authors would like to thank: Jeanne Anderson; David Autor; Kate Bartlett, Regina Barzilay; Joseph Blocher; Bernie Burke; Randy Davis; Ron Dolin; Stuart Elliott; Bruce Elvin; Kathleen Engle; Bill Freeman; Ron Friedman; Eddie Hartman; Nathalie Hofman; Tommi Jaakola; Jim Jones; Jonathan Kash; Boris Katz; Scott Krowitz; Maggie Lemos, Matt Levitt; Alex Levy; Marin Levy; Ruddick Lawrence; Matthew McCubbins; Kyla Moran; Jean O'Grady; Ross Pascal; Christina Patterson, Joe Propcopio; Bob Rowe; Craig Smith; Robin Toone; Doug Ventola; Jim Wagner; Ed Walters; Brad Wendel; Keith Winstein; Pat Ziegler; Yuan Zhang; and all participants of the 2105 Stanford Legal Ethics Schmooze, the William & Mary Faculty Law School Workshop, the Wake Forest Law School Faculty Workshop, and the MIT CSAIL/Economics lunch series. We are particularly grateful to Huron Legal's Sky Analytics for their generous provision of data and to the Spencer Foundation for supporting Frank Levy's work.

described how computers, capable of identifying relevant words and phrases, were displacing large numbers of lawyers in discovery practice. The article posed a warning to lawyers as it sought to make a broader point: computers could replace humans in a highly educated, white-collar occupation.

The warning is now common wisdom. Richard and Daniel Susskind argue that lawyers, among other professionals, face a future in which "increasingly capable machines, autonomously or with non-specialist users, will take on many of the tasks that are currently the realm of the professions."[2] Law professors John McGinnis and Russ Pearce contend that "the disruptive effect of machine intelligence" will "trigger the end of lawyers' monopoly."[3] Other commentators predict that "[i]n the not-too-distant future, artificial intelligence systems will have the ability to reduce answering a legal question to the simplicity of performing a search,"[4] and that "[o]nce we have fully artificial intelligence enhanced programs like LegalZoom, there will be no need for lawyers, aside from the highly specialized and expensive large-law-firm variety."[5]

---

[1] John Markoff, *Armies of Expensive Lawyers, Replaced by Cheaper Software*, N.Y. TIMES, Mar. 5, 2011, at A1.

[2] RICHARD SUSSKIND & DANIEL SUSSKIND, THE FUTURE OF THE PROFESSIONS, HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS 231 (Oxford Univ. Press 2015). *See also* RICHARD SUSSKIND, TOMORROW'S LAWYERS: AN INTRODUCTION TO YOUR FUTURE (Oxford Univ. Press 2013); RICHARD SUSSKIND, THE END OF LAWYERS: RETHINKING THE NATURE OF LEGAL SERVICES (Oxford Univ. Press 2010).

[3] John McGinnis & Russell Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041, 3066 (2014).

[4] Josh Blackman, *The Path of Big Data and the Law, in* BIG DATA AND THE LAW 5 (Kumar Jayasuriya ed., 2014).

[5] Paul Lippe & Daniel Martin Katz, *10 predictions about how IBM's Watson will impact the legal profession,* AMERICAN BAR ASSOCIATION (Oct. 2, 2014), http://www.abajournal.com/legalrebels/article/10_predictions_about_how_ibms_watson_will_impact (last visited Oct. 1, 2015); *Professor Dr Robot QC, Once regarded as safe havens, the professions are now in the eye of the storm*, ECONOMIST (Oct 17th 2015) at http://www.economist.com/news/business/21674779-once-regarded-safe-havens-professions-are-now-eye-storm-professor-dr-robot (last visited Oct. 20, 2015); Debra Cassens Weiss, *Will newbie associates be replaced by Watson?*, ABA JOURNAL (OCT. 26, 2015) at http://www.abajournal.com/news/article/will_associates_be_replaced_by_watson_computing_35_percent_of_law_firm_lead/?utm_source=maestro&utm_medium=email&utm_campaign=tech_monthly (last visited October 30, 2015). *See also* MARTIN FORD, RISE OF THE ROBOTS: TECHNOLOGY AND THE THREAT OF A JOBLESS FUTURE (Basic Books 2015); JERRY KAPLAN, HUMANS NEED NOT APPLY: A GUIDE TO WEALTH AND WORK IN THE AGE OF ARTIFICIAL INTELLIGENCE (Yale Univ. Press 2015).

Proponents of these arguments cite specific examples of computers performing lawyers' jobs. Predictive coding, the subject of Markoff's article, is a machine learning application that automates document classification in discovery practice.[6] Ross Intelligence, a legal application of IBM's Watson, advertises the ability to provide concise answers to natural language legal questions.[7] LegalZoom, RocketLawyer, and other online legal service providers produce basic wills, divorce agreements, contracts and incorporation papers without a lawyer's involvement.[8]

These technologies challenge the traditionalist view that lawyering is irreducibly human and force us to recognize that computers are changing the way law is practiced. Alone, however, they do not prove the imminent and widespread displacement of lawyers by computers. Much of the existing literature jumps to this conclusion and in doing so, foregoes the opportunity for a more nuanced and careful analysis.

In particular, the existing popular and scholarly literature suffers from three core weaknesses, which we seek to address. First, it fails to engage with technical details. We appreciate why—specifics blur the headlines and may be uninteresting to lay readers. But the details are critical for understanding the kinds of lawyering tasks that computers can and cannot perform. The details explain, for example, why document review in discovery practice is more amenable to automation than in corporate due diligence work, and why the automation of

---

[6] *See* Markoff, *supra* note XX.
[7] *See, e.g.,* Weiss, *supra* note XX. Ross.com is developing one such application, which it describes as your "brand new Super Intelligent Attorney." *See* ROSS, http://www.rossintelligence.com/ (last visited Oct. 20, 2015) ("You ask your questions in plain English, as you would a colleague, and ROSS then reads through the entire body of law and returns a cited answer and topical readings from legislation, case law and secondary sources to get you up-to-speed quickly.").
[8] *See, e.g.,* BENJAMIN BARTON, GLASS HALF FULL: THE DECLINE AND REBIRTH OF THE LEGAL PROFESSION (Oxford Univ. Press 2015); Benjamin Barton, *The Lawyer's Monopoly: What Goes and What Stays*, 82 FORDHAM L. REV. 3068 (2014); Benjamin Barton, *Lessons From the Rise of LegalZoom*, BLOOMBERG BNA (June 18, 2015), https://bol.bna.com/lessons-from-the-rise-of-legalzoom/.

Associated Press sports stories does not suggest the imminent automation of legal brief-writing.[9] We therefore offer a detailed review of salient legal technologies based on a set of unstructured interviews with computer scientists, legal technology developers, and practicing lawyers. We anchor our review in the present and near-term future since descriptions of artificial intelligence in a more distant future defy either proof or refutation.[10]

Second, existing work is unmoored from data on how lawyers spend and bill their time. For example, scholars suggest that the automation of document review is displacing large numbers of junior associates without reference to the amount of time junior associates previously spent on document review. Absent such data, their conclusions remain mere speculation. We seek to offer more reliable employment predictions, grounded in lawyer time usage data provided by Huron Legal's consulting arm, Sky Analytics.[11]

Finally, the existing literature fails to take seriously the values, ideals, and challenges of legal professionalism. Most scholars maintain that "professionalism" is mere cover for lawyer protectionism, and that the public interest is best served by commoditizing and computerizing as many legal services as possible.[12] Doing so, they contend, will lower costs and increase access.[13]

---

[9] *See infra* notes XX and accompanying text.

[10] For example, Susskind and Susskind argue the post-professional society will be reached "in the fully fledged, technology-based Internet society." SUSSKIND & SUSSKIND, *supra* note XX, at 232.

[11] Sky Analytics assists corporate clients in monitoring and analyzing the clients' legal expenditures. *See* SKY ANALYTICS, http://www.skyanalytics.com/ (last visited Oct. 20, 2015).

[12] Critics have long maintained that the professional form inures to the benefit of lawyers themselves at the expense of society by facilitating protectionism. *See, e.g.*, ANDREW ABBOTT, THE SYSTEM OF PROFESSIONS: AN ESSAY ON THE DIVISION OF EXPERT LABOR 184-86 (1988); RICHARD L. ABEL, AMERICAN LAWYERS 20 (1989); JEROLD S. AUERBACH, UNEQUAL JUSTICE: LAWYERS AND SOCIAL CHANGE IN MODERN AMERICA 88, 92, 99-102 (1976); MAGALI SARFATTI LARSON, THE RISE OF PROFESSIONALISM: A SOCIOLOGICAL ANALYSIS xvii (1977); SUSSKIND & SUSSKIND, *supra* note XX, at ch. 1. Certainly, individual lawyers and the organized bar have at times acted in protectionist and self-serving ways. But to conclude that the professional form is therefore undesirable is to ignore the many ways in which the profession is constitutive of the state and critical to the rule of law. *See* Robert W. Gordon & William H. Simon, *The Redemption of Professionalism?* in LAWYERS' IDEALS/LAWYERS' PRACTICES: TRANSFORMATIONS IN THE AMERICAN LEGAL PROFESSION 235 (Robert L. Nelson et al., eds.) (1992); *see also* Dana A. Remus, *Reconstructing Professionalism* (April 30, 2015), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2676094.

Access is undoubtedly a central challenge and goal of the legal profession, but computerization will not inevitably lower costs in all areas.[14] Moreover, adoption of new technologies implicates a range of additional and sometimes competing professional values that must be considered in a meaningful normative inquiry. We address these values, which include responsiveness to clients, compliance with law, creativity in legal argument, and democratic participation in law making.

Our focus is recent developments in legal automation, but we take as a given that earlier innovations dramatically impacted legal practice. Word processing revolutionized document drafting. The Internet permitted rapid document transmission and video conferencing; accelerated the breakdown of law firms' information monopoly on rates, services, and clients[15]; and increased clients' ability to spread legal work among multiple law firms.[16] Email increased the speed and ease of communication both among lawyers and between lawyers and clients, and expanded the number of associates a single partner could supervise.[17] These innovations changed law practice in fundamental ways. The next wave of technologies, our focus in this paper, promises similarly far-reaching effects.

---

[13] *See, e.g.,* John O. McGinnis and Russell G. Pearce, *The Great Disruption: How Machines Intelligence will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041, 3054-55 (2014); James E. Cabral et al., *Using Technology to Enhance Access to Justice*, 26 HARV. J. L. & TECH. 241, 246-56 (2012); William E. Hornsby, Jr., *Gaming the System: Approaching 100% Access to Legal Services through Online Games*, 88 CHI.-KENT L. REV. 917, 931-34 (2013); Ronald Staudt, *All the Wild Possibilities: Technology That Attacks Barriers to Access to Justice*, 42 LOY. L.A. L. REV. 1128-34 (2009); Michael J. Wolf, *Collaborative Technology Improves Access to Justice*, 15 N.Y.U. J. LEGIS. & PUB. POL'Y 759, 773-85 (2012).

[14] Many vendors are already patenting their new legal technologies, which will increase rather than decrease costs. *See* Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 IOWA L. REV. 1691, 1714 (2014).

[15] James Jones, The Changing Law Firm Risk Environment (full date) (unpublished manuscript) (on file with the Georgetown Law School Center for the Study of the Legal Profession).

[16] STEPHANIE KIMBRO, LIMITED SCOPE LEGAL SERVICES (A.B.A. 2012).

[17] Luis Garciano & Thomas Hubbard, *Earnings Inequality and Coordination Costs: Evidence from U.S. Law Firms* (Nat'l Bureau of Econ. Research, Working Paper No. 14741, Feb. 2009). Garciano & Hubbard speculate that the expanded capacity to supervise has, in turn, increased the earnings gap between senior partners and younger associates. *Id.*

Our discussion proceeds in three parts. In Part I, we address the current capabilities of computer technology in the legal realm. After reviewing basic aspects of machine intelligence, we discuss the potential for current or near-term automation of major categories of lawyers' work.

In Part II, we use data from Huron Legal's Sky Analytics to test two pieces of conventional wisdom regarding the impact of computers on lawyers—that the overall employment effects are significant, and that they are the greatest among junior associates. We show that there is no strong relationship between computers' employment effects and position within a firm, and that even where automation has made significant progress, its impact has been less than the headlines would have us believe.

Although computers are not displacing lawyer labor as rapidly as is often asserted, they certainly are changing the way law is practiced with significant and potentially unintended consequences for clients, our legal system, and the law itself. In Part III, we argue that the only way to understand and evaluate these consequences is to engage with the ways in which computers perform various lawyering tasks differently than humans. Sometimes, the differences will be beneficial by, for example, eliminating human error and increasing accuracy. Sometimes, however, the differences will be detrimental. They may create hidden risks of new errors or threaten creativity and flexibility in the law. These differences inform the desirability of automating various aspects of legal practice. They also shed useful light on the core values of legal professionalism and, ultimately, on the nature of law itself.

I. MACHINE COMPLEXITY, TASK COMPLEXITY, AND AUTOMATION

A set of basic ideas in artificial intelligence is a necessary starting point for any discussion of technology's impact in the legal arena.[18] Accordingly, we begin this Part with a selected overview of artificial intelligence capacities and limitations. We then review the current capacity of computer software to automate various lawyering tasks with reference to technologies that figure prominently in popular and scholarly writing on legal automation, including predictive coding, online legal service providers, automated news reporting, and IBM Watson's legal application.[19]

A. Some Basics of Machine Intelligence

We start with two observations: First, virtually all of a lawyer's tasks involve processing information,[20] and second, a computer processes information by executing instructions. It follows that for a computer to automate a lawyer's work, it must be possible to model the lawyer's information processing in a set of instructions.

The instructions that comprise a computer's model fall into two groups, which we refer to as *deductive* and *data driven* instructions. Deductive instructions can model tasks for which the

---

[18] This description of elements of artificial intelligence grows out of discussions over the last twelve years with MIT computer science faculty, in particular: Randall Davis, Peter Szolovits, and the late Seth Teller. For an earlier summary of this description, *see* FRANK LEVY & RICHARD J. MURNANE, THE NEW DIVISION OF LABOR, HOW COMPUTERS ARE CREATING THE NEW JOB MARKET (Princeton U. Press, 2005). *See also* TOMMI JAAKKOLA & REGINA BARZILAY, INTRODUCTION TO MACHINE LEARNING (manuscript, MIT 2015).

[19] In doing so, we employ the convention of the economics literature of describing a job or occupation (here, lawyering) as a set of tasks, so as to accommodate the fact that computers often replace parts rather than the entirety of a job. *See* David Autor, Frank Levy & Richard J. Murnane, *The Skill Content of Recent Technical Change: An Empirical Investigation*, 118 QUARTERLY J. ECON. 1279 (2003). For example, the ATM replaced part of a bank teller's job. *See* David Autor, Frank Levy, & Richard Murnane, *Upstairs, Downstairs: Computers and Skills on Two Floors of a Large Bank*, 55 Indust. & Lab. Rel. Rev. 432 (2002). With respect to law practice, it is necessary to include the automation of tasks previously performed by clerical staff and paralegals given that many legal technologies perform tasks that were not performed by lawyers in the first place.

[20] For example, a lawyer processes information about family relationships and assets into a will or transaction information into a contract. Information processing is central to virtually all human work.

structure of information processing is apparent[21]–an airline check-in kiosk, for example, which processes information from a credit card and the airline's reservation database into a boarding pass. The deductive instructions for the kiosk might read, in part:

> Does the name on the credit card match a name in the reservation database?
>
> If Yes: Check if the customer has a seat assignment.
>
> If No: Instruct the customer to see desk agent.

Where the structure of information processing is not readily apparent, it will be difficult if not impossible to model the task with deductive instructions. It may possible, however, to model the task with data driven instructions. Consider a mortgage underwriter trying to model a prediction of whether a mortgage applicant, if approved, will default within four years. The underwriter knows the individual's decision to default is related to information contained in the individual's application, including income, credit history, liquid reserves and the property's loan-to-value relationship, but the relationship is unlikely to be accurately modeled in deductive instructions. The underwriter can model the relationship in data driven instructions, however, by collecting a set of approved mortgage applications that are at least four years old and for which the underwriter can obtain information about whether the applicant defaulted within four years of receiving the mortgage. In the language of machine learning, each category of information that can be extracted from the applications is a feature of the data set. Statistical software is then used to estimate a model of mortgage default based on the values of these features.[22] The model

[21] In the early days of artificial intelligence, it was assumed most tasks could be described in deductive instructions (also called rules-based logic), but this has proven incorrect.

[22] Note that the statistical model does not attempt to describe the cognitive processing of the individual's default decision. Rather, it attempts to accurately predict the individual's default decision based on application information while treating the detailed decision process as a black box.

may involve extensive non-linearities,[23] but for simplicity, we represent it here as a single regression-like equation:

Equation 1: $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$

Where: $Y_i = 1$ if the i'th mortgage holder defaults

$= 0$ if the i'th mortgage holder does not default

X_{1i}, X_{2i} ….are the features drawn from the i'th mortgage holder's application

β_1 $\beta_{2\ldots\ldots}$ are the estimated coefficients for the application features

μ_i is a stochastic error term for the i'th mortgage holder

The estimation process is called "supervised learning"[24]: supervised because the parameters are required to align with whether the individual did or did not default; learning because the estimation process can be seen as learning the relationship (summarized in β's) between individuals' application information and their default decisions.[25] Once estimated, Equation (1) is a data driven instruction—a model that can be applied to features of new applications to predict the individual's *ex ante* default probability. Later in this paper, we discuss how similar modeling can be used to predict a judge's decision in a potential case.

Deductive instructions describe each step of a task's information processing task exactly (the airline check-in kiosk) and can predict the appropriate answer (the information on the

[23] Frequently, it will entail a neural net. *See* JAAKKOLA & BARZILAY, *supra* note XX. A complete model will also incorporate information from unsuccessful mortgage applications to adjust for "selection bias"—the fact that applicants who receive a mortgage are a non-random fraction of all applicants.

[24] A related technique, "unsupervised learning," is typically used to explore the nature of the sample rather than to estimate a model. This technique also begins with a sample of histories of successful mortgage applicants including whether or not they defaulted. The software estimates a set of coefficients for application features that divide the applicants into two groups or "statistical clusters" such that mathematical dissimilarity between the two clusters is maximized. In this process, default is treated as one more feature. For example, mathematical dissimilarity might be maximized by separating older Midwest applicants from younger Southeast applicants but the two clusters might have similar default probabilities.

[25] The estimation process is also described as training or as a form of pattern recognition, as the computer searches for the pattern of party and case characteristics that best predicts the judge's decision.

boarding pass) with perfect accuracy. Data driven instructions like Equation 1 merely estimate a relationship between the individual's default decision and application features and may therefore predict the individual's decision with error (the μ_i term). The model can be re-estimated over time, as individual mortgage histories accumulate and those histories are added to the training sample. But this will only lead to greater accuracy if an individual's decision is actually correlated with features for which the model accounts. If, in contrast, default decisions were driven by unanticipated shifts in the economy—a feature that did not appear in the model—the estimation would produce a set of statistically insignificant coefficients (β's) and the model as a whole would have no predictive power. Correspondingly, the estimate of Equation 1 also produces goodness-of-fit measures that summarize how well the model actually fits the sample—in the case of a regression, the measure is the squared correlation coefficient (R^2). Goodness-of-fit measures can be translated into the probability that the model's predictions involving current applicants are accurate.

Notwithstanding the capacity of data-driven instructions to handle a broad range of information processing tasks, some tasks are simply too complex or too opaque to be modeled for computers, using either deductive or data driven rules. Unstructured human interaction falls into this category because it frequently depends on formulating responses to unanticipated questions and statements, which, in turn, requires recognizing the context in which words are being used.

Frequently, recognizing context requires recognizing the affect of the person making the statement. Since Rossalind Pickard first coined the term "affective computing" in 1995,[26] significant progress has been made in enabling computers to recognize a user's affect by

---

[26] *See* ROSSALIND PICKARD, AFFECTIVE COMPUTING (MIT Univ. Press 2000).

measuring physiological states and facial expressions.[27] But as Pickard explains, it is one thing to differentiate between "user is frustrated" and "user is not frustrated," or even to differentiate between basic emotional states such as anger, fear, sadness, and love.[28] It is quite another thing, and "too much to ask of computers"[29] (at least at present), to recognize and label the infinite array of more complex emotional states that we ourselves can rarely label, but that we nevertheless navigate using the tacit skills of emotional intelligence.[30]

Returning to those tasks that *can* be expressed using deductive or data-driven instructions, an ideal model will specify an action for every potential contingency. It follows that automation is most successful in addressing structured tasks—tasks with a manageable number of predictable contingencies. For example, tech start-ups including Automated Insights and Narrative Science have received substantial publicity for developing models that write Associated Press-style summaries of baseball games. As we discuss below, this process is made possible because baseball is a highly structured game.[31] Similarly, the model of the default decision assumes a stable relationship between application features and the individual's decision to default (the set of constant β's) rather than a relationship that changes from meal to meal.[32]

When a task is less structured, as many tasks are, it will often be impossible to anticipate all possible contingencies. It may still be possible to automate the task, however, using one of three approaches. The first is to narrow its scope. If you swipe a credit card with a name that does not appear in the database of an airline check-in kiosk, the screen will read: "please see

[27] Juan Martinez-Miranda & Arantza Aldea, *Emotions in human and artificial intelligence*, 21 COMPUTERS IN HUM. BEHAV. 323 (2005); *Special Issue on Affective Computing*, 59 INT. J. HUM.-COMPUTER STUD. 1 (2003).
[28] Rossalind Pickard, *Affective Computing: Challenges*, 59 INT. J. HUMAN-COMPUTER STUDIES 57-58 (2003).
[29] *Id*.
[30] *Id.*
[31] *See infra* notes XX-XX and accompanying text.
[32] As we noted, however, that decision may not be correct.

desk agent." If your answer to an online will template's question: "Do you have children?" is anything but a simple yes or no, it will instruct you to consult a lawyer.

A variation on this approach is illustrated by Siri's response to some questions. If, on August 14, 2015 you had asked Siri: "Can a dog jump over a house?" it would have responded: "I'm sorry but I don't know the answer to that question." The check-in kiosk and the will template restrict the scope of a task using deductive instructions. Siri, in contrast, relies on data-driven instructions that look for similarities between a user's question and potential answers. The failure to produce an answer (the response: "I'm sorry but….") indicates the system's estimate that no potential answer has a sufficiently high probability of being correct.

A second approach to the problems of unanticipated contingencies is to minimize their likelihood by redesigning a task while maintaining its scope. In the 1970s, AT&T introduced automated speech recognition software to ask the recipient of a long distance collect call whether the recipient would accept charges. The original software, which imposed no structure on the recipient's reply, elicited many replies the software could not process: "That bum, I wouldn't speak to him…" AT&T quickly revised the software to ask "Will you accept charges? Please say 'yes' or 'no.'" Similarly, Amazon sidestepped the difficulty of automating the processing of free text email orders, with frequent misspellings or mistakes, by allowing a customer to order a book by clicking on an icon. In the legal context, early versions of Westlaw and Lexis overcame the inability of computers to process concepts by allowing users to conduct legal research by looking for particular words or combinations of words, rather than by area of law or precedential links. In the AT&T and Amazon examples, little was lost in the redesign. In the legal research example, in contrast, reducing the task to a key-word search distorted the results, leading Westlaw and Lexis to reintroduce broader approaches to legal research, such as through headnotes and indexes. In other cases, redesigning a task to reduce unanticipated contingencies may distort it so much as to make the results useless.

Pursuant to the third and most problematic approach, the computer ignores the unanticipated contingency and risks an incorrect result. Siri, like most software, is a work in progress. Two years ago,[33] it responded to the question above ("Can a dog could jump over a house?") with a list of kennels near the location where the question was being asked. Instead of recognizing the question as an unanticipated contingency that the software could not accommodate, it offered the best answer it could, albeit a nonresponsive one. Similarly, if the training data for the mortgage default model described above did not include an individual's alimony payments, the model may make poor predictions in situations where alimony payments cause an individual to default. In such situations, the model will not alert the user to the potential error because it will not have recognized the existence a problem.

To summarize before proceeding: tasks can only be automated if they can be modeled using deductive or data driven instructions. The relevant instructions must either specify an action for every contingency or adopt one of the three approaches just described. Highly complex and opaque information processing tasks, such as unstructured human interaction, cannot currently be modeled in this way.

B. Machine Intelligence in Legal Practice

With a basic understanding of the challenges and capacity of artificial intelligence as background, we turn to the potential for current or near-term automation of six categories of lawyering tasks. The tasks we address, which comprise the vast majority of legal practice, will strike legal readers as familiar. For non-legal readers, we recommend a brief look ahead to Table

[33] Specifically, March 9, 2013,

1 on page 33, which lists the range of tasks that lawyers routinely perform and percentages of time allocated to each.

The relevant software can be grouped in three categories based on the differing ways in which it affects lawyer employment. The first group automates all, or almost all, of a task previously performed by a lawyer. For example, once predictive coding software is trained, it will classify each of a set of documents as responsive or non-responsive, a task formally performed by junior associates or contract attorneys. The second group of software automates part of a lawyer's task. For example, a model that predicts how a judge would decide a case automates one aspect of legal strategy and analysis while other aspects remain unaffected. This partial replacement nevertheless reduces the number of lawyers required for a given volume of work. A third group of software, which we refer to as productivity software, automates heavily structured clerical tasks such as filing and calendaring. Prior to automation, these tasks were largely performed by clerical workers and paralegals such that automation has minimal impact on lawyers' employment per se. Below, we describe these categories of software as having "Strong," "Medium," and "Light" employment effects on lawyers. First, we review six categories of lawyering tasks and software addressing each.

1. Document and Case Management

We begin with document and case management, a set of tasks that run the gamut of machine complexity from quite easy to automate to very difficult to automate. On one end of the spectrum, document management has been successfully automated for years by networked computers and servers, and software that can sort and search files. New products have expanded to include automated templating, entry and billing of lawyers' hours, tracking of trust accounts, and related functions.[34] These are productivity applications, usually written in deductive instructions.[35] They replace tasks that used to be performed primarily by clerical personnel.

On the other end of the spectrum, in the related area of contract management, new data-driven applications are automating far less structured tasks.[36] KM Standards advertises software that reviews all of a company's contracts in a particular area, extracts the common provisions (apparently through unsupervised learning), and creates a basic template. The software then highlights discrepancies between the template and contracts proposed by other parties—tasks previously performed by associates. The technique is most effective when a company has a significant number of contracts in a particular area to ensure basic similarity in structure and language.

A third grouping of administrative work, case management, combines structured tasks such as billing and docketing with typically unstructured tasks such as assigning work to particular lawyers and supervising their performance. Many of the structured tasks have been successfully automated by the umbrella productivity applications described above and, as noted, primarily affect the employment of paralegals and legal assistants.[37] The unstructured tasks, in contrast, lie beyond the current capacity of computers. Tasks such as monitoring junior lawyers' work and dealing with parties who fail to honor contractual obligations require unstructured human interaction of a kind that computers cannot replace.

---

[34] *See, e.g.*, EFILECABINET, https://try.efilecabinet.com/capterra/ (last visited Oct. 22, 2015); LAWBASE, http://lawbase.com/ (last visited Oct. 22, 2015); ROCKET MATTER, https://www.rocketmatter.com/legal-billing-software/ (last visited Oct. 22, 2015); AMICUS, www.amicus.com (last visited Oct. 22, 2015); MYCASE, www.mycase.com (last visited Oct. 22, 2015).

[35] As we discuss below, many tasks automated by productivity applications were previously performed by clerical staff and paralegals such that any employment impacts are likely to fall on these two groups rather than lawyers.

[36] *See, e.g.,* ContractAssistant, www.contractassistant.com (last visited Oct. 22, 2015). This and other applications encompass such tasks as filing documents, identifying differences between successive drafts of contracts, and issuing alerts on due dates of contractual obligations.

[37] *See supra* note XX.

2. Document Review

Lawyers have been automating aspects of document review since the 1990s, when the explosion of electronically stored information created an overbearing volume of documents.[38] Keyword searching, programmed using deductive instructions,[39] was used to cull through massive document sets to find words or combinations of words that suggested relevancy.[40] Recall that keyword searching is an example of the second approach described earlier: a complicated task (document classification) is redesigned for a computer to avoid unanticipated contingencies. As is sometimes the case, the redesigned task produces results of limited utility. Because particular words do not necessarily correlate with particular meanings and content, keyword searching is often both under-inclusive (risking that important documents were being overlooked) and over-inclusive (raising the costs of review by returning large quantities of non-responsive documents)—problems with which many will be familiar from keyword searching on Westlaw or Lexis.[41]

In the late 1990s and early 2000s, recognizing that document review involved human information processing too opaque to be programmed using deductive instructions, programmers turned to data driven instructions.[42] The result was predictive coding, a machine learning application that, once trained, can replace human review. Similar to the mortgage default model above and pursuant to a standard protocol, supervising lawyers[43] review a "training sample" of documents (perhaps, 500) from among the full data set (likely totaling hundreds of thousands or

[38] Remus, *Predictive Coding, supra* note XX, at 1698.
[39] A typical keyword search rule involving the competitive behavior of a corporation might be: Select Document if: [Price] is within 15 words of ["customer"] or ["competitor"]. The program would then return all documents meeting the search criteria.
[40] Remus, *Predictive Coding, supra* note XX, at 1698.
[41] *Id.*
[42] *Id.,* at 1701-02.
[43] Initially, the training sample was coded by partners or senior associates, familiar with the case. Anecdotally, the task has already been pushed down to more junior lawyers.

millions of documents), classifying each document as relevant or not.[44] The software then scans the 500 document sample and extracts a set of linguistic features from each document, including the frequencies of specific words and phrases, and relationships among words. These features serve as independent variables in a statistical model designed to predict whether the document has been classified as relevant or not.

The supervising lawyers can test the statistical model by using it to classify a second sample of documents. If documents in the second sample are misclassified, they can add the misclassified documents to the original 500 document sample set and re-estimate the instruction. This iterative process continues until the lawyers are satisfied with the results. The estimated instruction model is then used to classify each of the remaining documents pursuant to an *ex ante* probability of relevancy. The lawyers choose a probability threshold such that documents with *ex ante* probabilities above the threshold are deemed as relevant.

Since 2012, when a federal judge first issued an opinion blessing predictive coding as an acceptable means of meeting discovery obligations,[45] the software has been used increasingly frequently.[46] Significant limitations remain, however. Experienced attorneys must still classify the training sample of documents and train the system's parameters, leading to up-front costs that make it efficient only for large cases that entail large volumes of documents. Moreover, while predictive coding is an effective means of reviewing documents as a part of discovery practice, it is far less effective in other contexts, including due diligence.[47] Document review in discovery is

---

[44] Remus, *Predictive Coding, supra* note XX, at 1702.
[45] *See* Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182, 185, 193 (S.D.N.Y. 2012), *adopted sub nom*. Da Silva Moore v. Publicis Groupe SA, No. 11 Civ. 1279(ACL)(AJP), 2012 WL 1446534 (S.D.N.Y. Apr. 26, 2012) ("Computer-assisted review now can be considered judicially-approved for use in appropriate cases.").
[46] *See, e.g.*, *Discovery,* HURON CONSULTING GROUP, http://www.huronconsultinggroup.com/Expertise/Legal/Discovery (last visited Oct. 21, 2015); *eDiscovery,* RECOMMIND, *http*://www.recommind.com/solutions/ediscovery (last visited Oct. 21, 2015).
[47] Telephone conversation with Nathalie Hofman, Huron Consulting (July 21, 2015).

a highly structured task: a single pattern of linguistics is used to classify an entire set of documents.

The due diligence that precedes a corporate transaction, in contrast, involves both structured and unstructured tasks. Firms are working to automate aspects of the document review of due diligence, but the task is far harder and progress therefore slower than in discovery practice. Apogee Legal[48] and Kira Systems[49] have developed software that crawls a company's network to identify, for example, vendor and sourcing contracts, customer agreements, software licenses, and leases (document types that occur in large volumes for which the software can be trained)[50]. The software is less effective in searching for a diverse set of documents where each type occurs in limited volume.

An equally important part of due diligence also involves searching for unexpected or surprising information and that may be located in a diverse set of documents—for example, potential conflicts under the Foreign Corrupt Practices Act. The task of searching for this information is far less structured and far more difficult to model accurately.

3. Document Preparation

The current capacity to automate document preparation hinges on the distinction between document drafting and legal writing. We define document drafting as producing legal documents that reflect the intent and agreement of the parties as accurately and unambiguously as possible. We define legal writing as producing written work that characterizes the state of the law and/or its application to a particular factual situation, either objectively or persuasively.

[48] APOGEE LEGAL, https://apogeelegal.com/ (last visited Oct. 21, 2015).
[49] KIRA, https://kirasystems.com/ (last visited Oct. 21, 2015).
[50] This is similar to KR Standard's creation of a corporation's standard contract template and similar limitations apply. *See supra* notes XX and accompanying text.

Document drafting is much more structured, and therefore easier to automate, than legal writing. Recent innovations build on lawyers' longstanding practice of creating and using basic templates for documents such as wills, trusts, and contracts. With the advent of personal computing, lawyers began saving the templates to personal desktops or document management systems. In a relatively modest step ahead, applications can now use deductive instructions to automate the process of field population.[51]

A more distinct innovation is the business model of online service providers like LegalZoom and Rocket Lawyer, which relies on the internet to market document templates directly to consumers. Legalzoom, for example, allows a consumer to obtain a number of legal documents from its website (including wills, powers of attorney, business filings, and bankruptcy or divorce petitions[52]) by indicating the document he or she is interested in and answering a series of document-specific questions. Based on the consumer's answers, the website produces a completed and customized document.

Although highly effective for basic and standardized legal documents, online document drafting is far more limited with respect to more complex and novel documents. For example, one commentator explains that template divorce agreements can be difficult to use because "uncontested divorces can quickly become contested divorces." The programs cannot anticipate all possible user situations, but only sometimes recognize and respond to unanticipated contingencies with the desired instruction to consult an attorney. At other times, they simply fail to anticipate a contingency, ignore it, and create an error.[53]

---

[51] Automated document drafting programs are frequently incorporated into document management software used by law firms. *See supra* note XX.

[52] *About Us*, LEGALZOOM, http://www.legalzoom.com/about_us.html (last visited Oct. 21, 2015).

[53] Peter Mahler, *LegalZoom LLC Agreement: Bargain or Blunder?* (June 15th, 2015) at http://www.nybusinessdivorce.com/2015/06/articles/llcs/legalzoom-llc-agreement-bargain-or-blunder/ (last accessed

Notwithstanding these limitations, automated document drafting has made significant progress in recent years. Legal writing, in contrast, has largely resisted automation. Commentators cite automated Associated Press summaries of baseball games and corporate earnings reports, noted above, to argue that this will soon change,[54] but the analogy does not hold. Extracting and summarizing relevant information about a baseball game or a company's reported earnings financial situation is a highly structured task. A baseball game can be largely reconstructed from the pitch-by-pitch game feed, and earnings reports have relatively structured and standard formats. To be sure, automating even these structured tasks entails substantial innovation.[55] But once the game (or the sentiment of the earnings report) has been reconstructed, writing the report involves a structured selection and listing of prewritten phrases with insertion of particular proper nouns (e.g. players' names). Note also that these short articles are usually directed at readers with limited information demands: a New York newspaper will contain an extensive, non-automated article on a Yankees' game while using Associated Press summaries to report on out-of-town games.

A legal argument expressed in a legal brief cannot be simplified in this way. Whereas a sports writer covers a settled game structure and a final definitive score, a lawyer writes amidst indeterminacy. Certainly, parts of a brief are standard and predictable—for example, the preliminary and concluding material, and the statement and explanation of relevant standards of review. But the articulation and explanation of an argument is the product of conceptual creativity and flexibility that computers cannot exhibit. The analysis section of a legal brief

---

Oct. 20, 2015); Gene DuBose, *A LegalZoom Horror Story*, at http://www.ezlaw.org/newsletters/LegalZoom%20Horror%20Story.pdf (last accessed Oct 20, 2015).

[54] McGinnis & Pearce, *supra* note XX, at 3051.

[55] One of the most successful companies, Automated Insights, supplements its game feed reconstruction with a running estimation of win/loss probabilities. When a particular play causes a significant shift in probabilities, the play receives heightened attention in the article.

requires a complex interplay between law and fact, in which the law that governs is determined by the facts while the relevant facts are determined by the governing law. It requires the use of precedent, which while second-nature for a lawyer, would be exceedingly difficult (currently impossible) to model for a computer. A single case can be used to support two opposing positions; arguing for one as opposed to the other requires an ability to contextualize the case in a line of precedent and to distinguish between binding holding and non-binding dicta. Often, an effective legal argument also requires the ability to transplant concepts from one area of law to another in order to argue for a novel legal theory or change in the law. These unstructured and opaque conceptual tasks lay far beyond the current capacity of computers.

4. Legal Research and Reasoning

Legal research was transformed by the 1970s emergence of computer-assisted research services such as Westlaw and Lexis. Commentators claim that we are on the brink of a second transformation with the emergence of IBM Watson based question and answer systems such as the one currently being developed and marketed by Ross Intelligence.[56] If IBM Watson can win Jeopardy, these commentators reason, surely it can answer legal questions.[57]

While computerized research and information retrieval tools are unquestionably advancing, we view these claims as premature. To understand why, it is useful to consider three core questions that form the core of information retrieval software (from a computational perspective):

- What form of answer is the user looking for – a factoid or a more complex concept?
- What database will the user be searching: the world-wide web or a data base tailored more narrowly to the query?

[56] ROSS, http://www.rossintelligence.com/ (last visited Oct. 20, 2015).
[57] *See e.g.,* Weiss, *supra* note XX.

- What mechanism is used to link a search query to potential answers?

With respect to the first question—the form of the answer—a few minutes working with Google or Siri demonstrates that questions requiring a factoid often result in the right answer while questions requiring a more complex answer typically return a collection of passages from various sources with various degrees of relevance. This difference reflects computers' current inability to summarize the conclusions of multiple passages. Quoting Daniel Jurafsky:

> Extractive summarization answers the query by pulling a set of short text pieces from the documents (snippets). Google, for example, responds to queries in this way. Abstractive summarization expresses the ideas at least in part in different words. Abstractive summarization is much closer to the language one would find in a legal memo and it is currently an important research goal, but very difficult.[58]

It follows that a lawyer can expect a computer to produce raw material for a legal summary, but will still have to write the summary itself. Equally important, the difficulty in automated abstractive summarization reflects a more general difficulty of modeling abstract concepts that rely on synonyms, hyponyms,[59] analogies (including novel applications of precedent), and other subtle uses of language some of which may not be in the computer's data base. We return to this point below.

The design of a database rests on competing considerations. If the data base is too narrow, it will exclude potential answers. But if the data base is too broad, it may return irrelevant matches (given that most information retrieval systems function via elaborations of word matches). Typical legal databases address this by giving users a choice—for example, to search cases in a particular district court to maximize precision or a search of all federal court

---

[58] *See* Daniel Jurafsky & Christopher Manning, *Natural Language Processing*, COURSERA, https://class.coursera.org/nlp/lecture/191 (last visited Oct. 22, 2015)

[59] The hyponym of a word or term is another word or term with a more specific meaning: a sedan is a hyponym of a car.

cases to maximize recall.[60] Most users balance the two goals by searching a number of different databases.

The third critical design question is the mechanism used to link a user's search query with a potential answer. A user of Westlaw and Lexis can always begin with a key word search, but keyword searching, as noted, is frequently both under and over-inclusive. Thus, the services offer two ways of limiting and narrowing results. A user can limit the searched database so as to examine only cases in a relevant jurisdiction or time period. Alternatively, a user can start with an indexing tool for a broader entry into a particular area of law.

Constructing these indexing tools involves a significant amount of human time and processing, and is therefore expensive. At the outset, humans write a summary headnote for each case filed in the system (reflecting the difficulty of automating abstract summarization). For Lexis, humans then use these headnotes to classify cases into the Lexis Topic system. For Westlaw, a machine learning algorithm links the head notes to the West Key typology codes.[61]

A different way of narrowing the results is through the ranking of results. Traditional versions of Lexis and Westlaw listed results by reverse chronological order or frequency of search terms. FastCase, a more recent entrant to the market, ranks results by relevancy, determined primarily by citation frequency combined with the relative importance of the citation.[62]

---

[60] Precision refers to the fraction of returned answers that are relevant. Recall refers to the fraction of all relevant answers that are returned. Generally speaking, the more focused the database, the higher the precision; the broader the database, the higher the recall.

[61] *See* Susan Nevelow Mart, *The Relevance of Results Generated by Human Indexing and Computer Algorithms: A Study of West's Headnotes and Key Numbers and LexisNexis's Headnotes and Topics*, 102 L. LIBR. J. 221 (2010).

[62] *See What is Fastcase?*, FASTCASE, http://www.fastcase.com/whatisfastcase/ (last visited Oct. 22, 2015). New versions of Lexis and Westlaw—Lexis Advance and WestlawNext—are incorporating relevancy rankings as well, based on a combination of features such as past search patterns, document characteristics, and matching terms.

A more fundamental departure from Westlaw and Lexis's approach to linking questions and answers comes from IBM Watson applications. For the foreseeable future, however, this approach requires similarly significant investments of human time and energy and is similarly expensive. At the outset, the project supervisor has to populate the project's database by legal documents that have been broken into paragraphs or passages. The supervisor appends a set of natural language practice questions to each paragraph, such that the paragraph is a good answer for each of the attached questions. Each practice question must be worded in multiple ways to reduce the likelihood that the software will fail to recognize a user's question as having the same meaning as a previously processed practice question.

In simplified description,[63] the software then processes each practice question and each paragraph into mathematical vectors of a common set of linguistic features,[64] where each feature is associated with a statistical coefficient (as in Equation (1)). Through a training process, these coefficients are adjusted to maximize the chances that the program will deem a paragraph's vector to be mathematically closest to the vectors of its associated practice questions. When a user enters a new question, the question is converted to a vector of the same set of linguistic features and the software defines the best answer as the paragraph whose feature vector is mathematically closest to the feature vector of the user's question. Thus in the diagram of Figure 1, Paragraph 1 is selected as a better answer than Paragraph 2 where mathematical closeness in this description is measured by the cosine of the angle between vectors. [65]

---

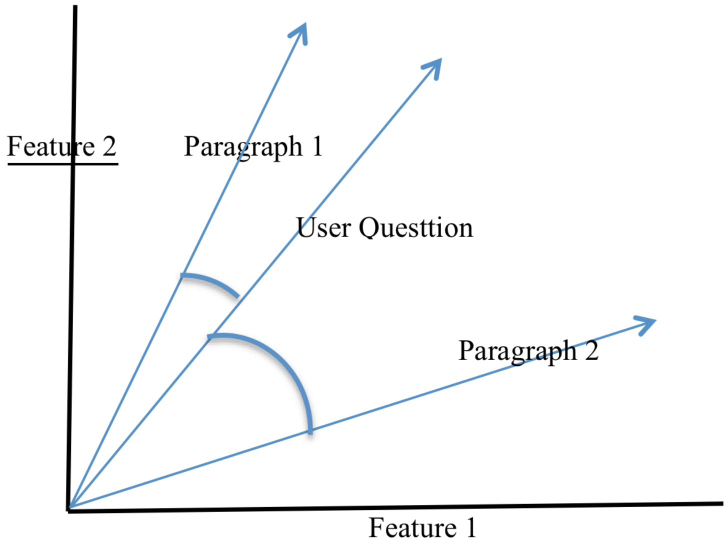
[63] For a more complete description of Watson's question answering architecture, *see* David Ferrucci et al,, *Building Watson: An Overview of the Deep QA Project,* 31 AI MAG 59 (2010).

[64] Features can include such items as word frequency, inverse word frequency and the presence of specific relationships among entities.

[65] The cosine similarity is one of several ways to measure mathematical closeness.

Figure 1
A Graphic Representation of Choosing the "Closest" Answer

(User Question is more Similar to Paragraph 1 than Paragraph 2)



This system relies on building statistical links between questions and answers—links ultimately based on the actual words in a question and answer as well synonyms and hyponyms that draw on word networks stored in the language processing software. These links are built through extensive training of the system in which the system receives feedback on whether a given answer is correct or incorrect. Once in use, however, a system can confront a potentially large number of potential questions, each of which can be expressed in a variety of ways (particularly if a question involves abstract concepts and analogies). Thus, it is not surprising that even an extensively trained system will confront questions for which it has not been adequately trained. As we have seen, the program can respond in one of two ways. It can state that it cannot answer the question (because the estimated probability of error is too high) or, more problematically, it can produce its "best" answer. If on November 27, 2015, Siri had been asked "Can a dog jump over a house?," the response would have been a link to a child's riddle about a dog jumping over a dog house and a second link to an ASPCA bulletin on teaching dogs not to jump.[66]

Because these systems perform best on questions on which they have been trained, a user is likely to receive better results with general, rather than highly detailed, inquiries. For example, a system will likely respond to the following question with a set of paragraphs, many of which will be relevant:

"When can a debtor reject a collective bargaining agreement?"

By contrast, a system will likely respond to the following, more detailed version of the question with some relevant paragraphs, but not all that are needed to construct a full answer:

> "Can a debtor reject a collective bargaining agreement where debtor is a city that filed for Chapter 9 bankruptcy and previously attempted to negotiate with a private union before rejecting its collective bargaining agreement?"

In considering the employment impacts of both search and question answering tools, it is important to recognize the substantial remaining human role in defining and directing legal research. Consider, for example, the nature of a case law search, which frequently begins with an initial set of controlling cases. As Susan Mart writes:

> [I]t is rare that the facts of those cases are so close to the facts of the client's case that your research is complete. The second part of the research project then begins—the search for case-specific relevant authority. The researcher needs to find other cases, similar in legal conclusions and more similar factually to the client's case. This search for more specifically relevant primary law can be called "level two research." The researcher uses the major and controlling cases in the relevant area of the law (however located) as seed documents to link forward through headnotes, key numbers, KeyCite, and Shepard's or backward through headnotes, key numbers, and the cases cited in the seed cases. This type of forward and backward searching from seed documents is instrumental for finding "application cases"—cases that have only marginal value as support for an abstract proposition of law, [but] have great value in their application of the proposition to facts similar or analogous to the facts of your own case.[67]

[66] We do not know why the system did not retain the "I don't know" response of three months earlier.
[67] Mart, *supra* note X, at 222.

Mart describes an iterative process in which a lawyer specifies the parameters for a search, which the software then performs. It is therefore the search that has been automated, not the entire task of researching precedents. A similar logic applies to question answering systems. They can automate an actual search, perhaps more effectively than Westlaw or Lexis, but they cannot automate the designation of search parameters. That work remains for lawyers—most often, for associates.

Beyond information retrieval, computers have made significant progress in one other area of legal research and reasoning—prediction. In recent years, software such as Ravel Law and Lex Machina have collected and analyzed massive amounts of data on judges and their decisions, producing data-driven statistical prediction models that are broadly similar to the supervised learning model of Equation (1) above.[68] Again, however, a significant human role remains in interpreting the data and formulating advice for clients, a topic we return to shortly.

5. Interpersonal communication and interaction

A fifth category of work encompasses a diverse range of tasks unified by a common characteristic—the centrality of interpersonal interactions. This category includes: communications (with clients, opponents, and others), client counseling, factual investigation, and negotiation.

Technology has affected lawyers' interactions in ways that allow for the automation of some of these tasks. Social media has opened new avenues of business development, and a

[68] For example, Ravel's website explains: "By analyzing millions of legal documents, Ravel provides strategic insight into an array of factors that affect a judge's decision-making." *See* RAVEL, https://www.ravellaw.com/ (last visited Oct. 22, 2015). *See also* LEX MACHINA, https://lexmachina.com/what-we-do/(last visited Oct. 22, 2015) ("We mine litigation data, revealing insights never before available about judges, lawyers, parties, and patents, culled from millions of pages of IP litigation information. We call these insights Legal Analytics®, because analytics involves the discovery and communication of meaningful patterns in data.").

handful of lawyering interactions, such as court filings and docketing, have been fully automated. The vast majority of a lawyer's personal interactions, however, continue to require spontaneity, unstructured communication, and emotional intelligence. Examples are plentiful: A lawyer may need to push a client to execute a will; spend hours interviewing a criminal defendant to develop enough trust to elicit full information; or read a deponent's facial expression and body language to determine how to proceed with questioning. Moreover, many individual clients report that a lawyer's trustworthiness and ability to provide a close and personal relationship are among the most important traits they look for from a lawyer—far more important than the lawyer's training, competence, or specialty.[69]

With respect to client advising, computers have made significant progress in one area just noted—prediction. For at least three reasons, however, most client advising remains outside of the current domain of automation. First, these programs address courts and case law. Lawyers must routinely predict many other things, such as how an opponent will reach to a settlement offer or how an agency will interpret a regulation. Second, many clients want more than a series of statistical probabilities. They want a lawyer's judgment and assurances as to what course of action will most effectively serve their short and long term interests. Some clients want this for their own comfort; others want it to reassure affected constituents; still others want it for

---

[69] *See* COREY S. SHDAIMAH, NEGOTIATING JUSTICE: PROGRESSIVE LAWYERING, LOW INCOME CLIENTS AND THE QUEST FOR SOCIAL JUSTICE (2009) (citing interviews of clients expressing that friendship and trust were at the forefront of what they wanted from lawyers); Marcus T. Boccaccini et al., *Client-Relations Skills in Effective Lawyering: Attitudes of Criminal Defense Attorneys and Experienced Clients*, 26 LAW & PSYCHOL. REV. 97, 111 (2002) (citing a poll in which clients ranked obtaining clients' opinions, spending time with clients before court, and keeping clients informed of their cases as among the things they cared most about in a lawyer; also citing evidence that inmates cared more about a lawyer who cared about them, would be honest, and would spend time with them before their court date than about the lawyer's skills); Marcus T. Boccaccini & Stanley L. Brodsky, *Characteristics of the Ideal Criminal Defense Attorney from the Client's Perspective: Empirical Findings and Implications for Legal Practice*, 25 LAW & PSYCHOL. REV. 81 (2001); Anne E. Thar, *What Do Clients Really Want? It's Time You Found Out*, 87 ILL. B.J. 331 (1999).

purposes of a potential advice of counsel defense.[70] Third and most importantly, effective advising encompasses more than prediction. It requires a lawyer to understand a client's situation, goals, and interests;[71] to think creatively about how best to serve those interests pursuant to law; and sometimes, to push back against a client's proposed course of action and counsel compliance.[72] These are things that frequently require human interaction and emotional intelligence[73] and cannot, at least for the time being, be automated.

Fact investigation, a task that may seem far afield, is similarly dependent on unstructured communications. Some aspects of the task can be automated. For example, software can usefully pull together vast amounts of online data regarding a client or opponent, and some lawyers and legal aid clinics automate initial client intake. For the most part, however, factual investigation is a highly unstructured task that resists automation. It frequently entails interviews in which significant amounts of information may be transmitted nonverbally, in ways a computer would have difficulty detecting. It also requires flexibility from a lawyer, beyond the capacity of a computer, in adjusting the relevant questions as new information is discovered.

A final task in this category—negotiation—often requires personal interaction and effective use of emotion. Negotiation theorists explain that skill in reading an opponent's emotions allows a negotiator to achieve greater understanding of the opponent's interests and

---

[70] Many clients may want a lawyer's advice as a means of avoiding what behavioral economists refer to as "regret"—the guilt and responsibility that can accompany a wrong decision in an uncertain situation. *Cf.* Richard H. Thaler*, Toward a Positive Theory of Consumer Choice,* 1 J. EC. BEHAVIOR & ORG 54 (1980) (observing that one reason doctors look for second opinions is to share responsibility and reduce regret for diagnoses that may turn out to be wrong).

[71] Katherine Kruse, *Beyond Cardboard Clients in Legal Ethics*, 23 GEORGETOWN J. LEGAL ETHICS 103, 104 (2010); Katherine R. Kruse, *The Promise of Client-Centered Professional Norms, Symposium: Restorative Justice and Attorney Discipline,* 12 NEV. L.J. 341, 346 (2012); Marc Lauritsen, *Liberty, Justice, and Legal Automata*, 88 CHI.-KENT L. REV. 945, 954 (2013); William H. Simon, The Ideology of Advocacy, 1978 WIS. L. REV. 29; 53.

[72] DEBORAH L. RHODE, IN THE INTERESTS OF JUSTICE (2000); Robert W. Gordon, *A New Role for Lawyers?; The Corporate Counselor After Enron*, 35 CONN. L. REV. 1185 (2003); Tanina Rostain*, Ethics Lost: Limitations of Current Approaches to Lawyer Regulation,* 71 S. CAL. L. REV. 1273, 1274-75 (1998); W. Bradley Wendel, *Professionalism as Interpretation*, 99 NW. U. L. REV. 1167, 1171-72 (2005).

[73] Remus, *Reconstructing Professionalism, supra* note XX at 25-29.

concerns, to assess risk more accurately, and to deploy negotiation tactics more effectively. Online dispute resolution programs are rendering these human skills unnecessary in a small but growing category of cases, however.[74] For example, a company called Modria markets online dispute resolution to companies.[75] Its website describes that it gathers relevant information regarding the dispute, summarizes areas of agreement and disagreement, and makes suggestions for resolving the issue.[76] It does so through deductive instructions, rendering negotiation (as lawyers understand the task) unnecessary.[77] Currently, the approach is used primarily for small ecommerce disputes, but Modria is expanding into larger and more complicated types of disputes.[78] A number of other companies are developing similar products,[79] while legal reform groups are encouraging courts to increase efficiency and manage dockets through use of such products.[80]

---

[74] *See* Richard Susskind and Matthew Levy, Likely Developments in ODR (Feb. 16, 2015), available at https://www.judiciary.gov.uk/publications/likely-developments-in-odr/; *see also* William E. Hornsby, Jr., *Gaming the System: Approaching 100% Access to Legal Services Through Online Games*, 88 CHI.-KENT L. REV. 917, 932 (2013).

[75] *About*, MODRIA, http://www.modria.com/about/ (last visited Oct. 20, 2015).

[76] *Id.*

[77] For example, a deductive instruction could read: ("If (Customer is Low Risk) and (Dispute Amount is less than $10) and (Customer Disputes Filed Account Lifetime is 0) then (Authorize Full Refund) and (Close Case)."). *Id.*

[78] Eric Johnson, *Modria Wants You To Settle Your Workplace Problems (and Even Patent Disputes) Online*, ALL THINGS D (Nov. 24, 2012), http://allthingsd.com/20121124/modria-wants-you-to-settle-your-workplace-problems-and-even-patent-disputes-online/ (last visited Oct. 20, 2015). *See also* Interview with Founder of Modria, at https://www.judiciary.gov.uk/publications/modria-the-operating-system-for-odr-video-extract/ ("[O]ur goal is to be the operating system for online dispute resolution. So any kind of dispute, no matter how complicated or how simple, how high volume or low volume, we can use these building blocks at Modria to build an appropriate resolution path for that dispute. So that's our objective.")

[79] See, e.g., Graham Ross and Beth Silver, Case Studies (Feb. 16, 2015) available at https://www.judiciary.gov.uk/wp-content/uploads/2015/02/Additional-Case-Studies1.pdf.

[80] *See, e.g.,* CJC ODR Advisory Group, Online Dispute Resolution for Low Value Civil Claims (February 2015) available at https://www.judiciary.gov.uk/reviews/online-dispute-resolution/odr-report-february-2015/ (recommending a new Internet-based court service, which would offer three services: Online Evaluation, which would help users to understand and evaluate their potential claims; Online Facilitation, which would facilitate early resolution of disputes without the involvement of a judge; and Online Judges, who would decide parts or all of cases through structured online pleading).

Additional new technologies are emerging to aid lawyers in negotiating by, for example, analyzing and representing the overlap between two parties' preferences.[81] Such programs address one or at most two issues, and their resolution is constrained by the parties' stated initial preferences. They nevertheless suggest that computers may eventually play a larger role in aiding, if not replacing, lawyers' negotiating work.

6. Courtroom Appearances

A final category of work, courtroom advocacy, is distinct from the others insofar as even the most fervent technology advocates are not predicting near-term automation. In part, this is because the policies and restrictions of unauthorized practice of law rules operate at their strongest in the courtroom. More fundamentally, it is because effective advocacy requires emotional engagement with the decision-maker.[82] As two experienced advocates explain:

> An inexperienced trial lawyer's dull and confusing closing argument in a complex business dispute will create negative feelings of boredom and frustration in the minds of the jurors…an accomplished advocate can communicate to the juror the facts of the identical dispute in a way that will evoke positive emotions about justice and fairness in the marketplace.[83]

It is not only in arguments to a jury that emotion is critical. The emotions a lawyer deploys to persuade a judge may differ from those designed to persuade a jury, but emotion is a critical spur to all action and decision-making.[84] And yet, the field of affective computing is nowhere near enabling computers to foster, recognize, and respond to the full range of human emotions.

---

[81] For example, Keith Winstein, Assistant Professor of Computer Science, Stanford University, has shown that some telecom related negotiations on access prices could be solved by an online auction. *See* http://cs.stanford.edu/~keithw/.

[82] *See, e.g.,* ARISTOTLE, RHETORIC (Pearson, 1st ed. 1960); RICHARD DU CANN, THE ART OF THE ADVOCATE 156 (Penguin Books 1980); EMOTION IN ADVOCACY (2003).

[83] John C. Shepherd & Jordan B. Cherrick, *Advocacy and Emotion*, 3 J. ASSOC. OF LEGAL WRITING DIRECTORS 152, 152 (2006) (originally published as 138 F.R..D. 619 (1991)). *See also* MOOY, ADVOCACY AND THE ART OF STORYTELLING 2 (1990).

[84] Shepherd & Cherrick, *supra* note XX, at 153.

* * *

To summarize, computer technology is the most advanced in the areas of document review and document management, though much of the work in these areas is performed by clerical staff and paralegals such that the technology will have limited employment impacts on lawyers. Computer technology is the least advanced in the areas of legal writing, advising clients, communications and interactions, factual investigation, negotiations, and court appearances. The tasks of case administration, due diligence, document drafting, legal research, and legal analysis and strategy fall somewhere between the two extremes.

II. EMPLOYMENT EFFECTS

We have seen that computers can automate structured and repetitive information processing tasks, but that they have much more difficulty with unstructured and opaque information processing. How does this map onto the actual work of lawyers? In this Part, we combine the foregoing discussion with data on lawyer time usage to test two frequently advanced claims regarding the employment impacts of legal automation: that automation is rapidly displacing lawyer labor, and that that this displacement is greatest among junior associate positions. Using illustrative calculations of automation's likely effects if all currently available technology is adopted, we show that neither claim is well grounded.

A. The Data

Our data on time usage comes from Huron Legal's Sky Analytics of Framingham, Massachusetts,[85] a consulting firm that provides corporate clients with aggregation and analysis of invoices billed by law firms. Typically, each invoice covers a small increment of time and describes the work the lawyer performed by reference to a task code from the ABA's Uniform

[85] SKY ANALYTICS, http://www.skyanalytics.com (last visited Oct. 24, 2015).

Task-Based Management System (UTBMS).[86] The UTBMS consists of 114 distinct task codes, which we have aggregated into 13 categories for purposes of identifying patterns.[87] Sky Analytics supplements the invoice with information on the submitting lawyer, including their status within the firm (associate or partner) and how many years they have been practicing.

We note several limitations of the data at the outset. The original UTBMS codes (and hence our 13 aggregated codes) allow lawyers significant discretion in how they record their time. Errors can entail both mislabeled time, and inaccurately recorded amounts of time. Billing partners may also revise time allocations prior to sending an invoice to the client. For example, one interviewee explained that clients do not like to see large amounts of time invoiced to legal research so billing lawyers might reallocate that time to the task the research is associated with, the broad category of "legal analysis and strategy," or a category of unbilled time.[88]

Moreover, because the invoices come from corporate clients, SkyAnalytics cannot provide a complete set of invoices billed by a single or several law firms. That said, given that the invoices Sky Analytics tabulated in 2014 totaled $2.31B, we (and Sky Analytics) believe that a pooled sample of all billing from firms with 1,000 or more lawyers (Tier 1 firms) provides a rough approximation of the distribution of hours billed to each task by junior associates (2 years or less), senior associates, and partners in a typical large law firm.[89] Notably, the data suggests that time-on-task among smaller sample firms (Tiers 2 through 5) follow a similar distribution.

[86] *See Uniform Task-Based Management System,* AMERICAN BAR ASSOCIATION, http://www.abanet.org/litigation/utbms/home.html (last visited Oct. 21, 2015) (a set of billing codes designed to standardize the categorization of tasks).
[87] Our thirteen aggregated tasks are: Advising Clients; Other Communications/Interactions; Case Administration and Management; Court Appearances; Document Drafting; Document Management; Document Review; Due Diligence; Fact Investigation; Legal Analysis and Strategy; Legal Research; Legal Writing; Negotiation.
[88] Telephone interview with Jean O'Grady, Author of the Dewey B Strategic Blog and Director of Research Services at an Amlaw100 law firm (July 22, 2015).
[89] The distribution of hours billed is not precisely the same as the distribution of tasks in the firm because firms bill less than 100% of junior associates' hours.

Finally, the data provides no information on the work patterns of solo practitioners, who comprise about 40 percent of all practicing lawyers,[90] or contract attorneys, whether hired by the law firm or the client.[91]

For purposes of this project, Sky Analytics provided us with a blinded data set for 2014 that allowed us to construct the following information:

- Distribution of hours billed by task.
- Distribution of hours billed by task further disaggregated by law firm size in five "Tiers" (Tier 1 > 1,000 lawyers through Tier 5 < 25 lawyers).
- Distribution in of hours billed by task and law firm size further disaggregated by position in the firm: (Associate ≤ 2 years; Associate > 2 years; Partner).

Table 1 lists the thirteen aggregated task codes with two distributions of hours spent on task: the 2014 distribution of time on task billed by all Tier 1 firms (> 1,000 lawyers) and the 2014 distribution of time on task for all Tier 2-5 firms (all other firms in the Sky Analytics Sample). Table 1 also indicates the employment effects on lawyers of each task, reflecting the discussion and conclusions of Part I.

---

[90] CLARA N. CARSON & JEEYOON PARK, THE LAWYER STATISTICAL REPORT: THE U.S. LEGAL PROFESSION IN 2005 6 (A.B.A. 2012).

[91] Nor does the data account for time billed to business development or other internal matters not billed to clients.

Table 1
Percent of Invoiced Hours Spent on Various Tasks,
Grouped by Estimated Extent of Computer Penetration

Task	Tier One Firms	Tiers Two – Five Firms
Strong Employment Effects	**4.1%**	**3.6%**
Document Review	4.1%	3.6%
Moderate Employment Effects	**39.7%**	**40.4%**
Case Administration and Management	3.7%	5.6%
Document Drafting	5.0%	4.0%
Due Diligence	2.0%	3.4%
Legal Research	0.5%	0.4%
Legal Analysis and Strategy	28.5%	27.0%
Light Employment Effects	**56.0%**	**55.7%**
Document Management	0.4%	0.7%
Fact Investigation	9.2%	9.6%
Legal Writing	11.4%	17.7%
Advising Clients	9.3%	3.2%
Other Communications/Interactions	8.8%	5.0%
Court Appearances and Preparation	13.9%	14.5%
Negotiation	3.0%	5.0%
Totals **	99.8%	99.7%

** Percentages may not sum to 100% due to rounding.

Note that distributions of time-on-task are similar for Tier 1 firms and Tier 2-5 firms. In particular, only 4.1 percent of lawyers' time at Tier 1 firms, and 3.6 percent of time at Tier 2-5 firms was billed to tasks where automation has potentially strong employment effects. One could argue that these low percentages reflect the impact that particular technologies have already had—most notably the impact of predictive coding in automating document review. However, predictive coding was not widely used until it was officially blessed by a federal judge in 2012, and Sky Analytics data shows that in 2012, lawyers at Tier 1 firms billed only 6 percent of their time to document review.

More likely, the low percentage is explained by two factors. First, document review in our typology covers only discovery practice, not due diligence (which, as described above, is far harder to automate). Accordingly, associates in departments other than litigation would not devote any of their time to the task. Second, clients have been pressuring law firms for over a decade to hold down litigation costs including outsourcing, offshoring, or using contract attorneys to perform document review. These pressures intensified following the 2008 financial collapse. Data from the U.S. Department of Commerce *National Income and Product Accounts* on the value of legal services show that between 1990 and 2007, the value of legal services (adjusted for inflation) grew at an average rate of 12.8 percent. Between 2007 and 2013, the growth rate was -0.6 percent while the number of lawyers in the country continued to grow (by 2.7 percent).[92] The shift in supply versus demand empowered clients, who could insist on cost-

[92] U.S. Bureau of Economic Analysis, *National Income and Product Accounts,* Tables 6.1 B-C, available at http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=1&isuri=1 (last accessed Nov. 30, 2015).

cutting measures, including outsourcing and the exclusion of junior associates from their matters. Thus, the task may already have been pushed out of domain of firm lawyers' work by 2012.[93]

B. An Illustrative Calculation

To develop a better sense of the likely employment effects if all currently existing legal technologies were fully adopted, we offer an illustrative calculation based on computers' employment effects in performing various lawyering tasks and the percentage of lawyer time spent on those tasks. We recognize that automation is only one among several factors that shape the market for legal services.[94] But here, we focus narrowly on automation's current impact by adopting the "lump of labor fallacy"—the assumption that there is a fixed amount of work to be done such that the automation of any task results in reduced employment. This assumption is wrong in the market for legal services, just as it is wrong elsewhere. By many estimates, more than 75 percent of civil legal need in the country goes unmet.[95] The automation of lawyering tasks may address this latent market rather than replacing existing lawyer labor. Alternatively, it may push lawyers to serve this latent market as a means of finding new work. Nevertheless, the assumption offers a transparent basis on which to estimate automation's effect. Our calculation also assumes that the quality of lawyers' work remains constant—that lawyers use technology to produce a constant product in less time rather than an improved product with no reduction in time.

We begin in Table 2 with the distribution of hours spent on tasks in large law firms (employment = 1,000 +). We recognize that large law firms employ only a small fraction of all

[93] *See, e.g.*, A. Jones & J. Palazzolo, *What's A First-Year Lawyer Worth?*, WALL. ST. J., October 17, 2011 (reporting that 20 percent of corporate legal departments insist that no first or second year attorneys not work on their matters).
[94] Consider, for example, that the introduction of Automatic Teller Machines (ATM) was expected to quickly reduce the number of tellers per 1,000 population. In reality, the ATM effect was offset for a time because many bank corporations began to compete by opening large numbers of branches.
[95] *See* Deborah L. Rhode, *Whatever Happened to Access to Justice?,* 42 LOY. L.A. L. REV. 869, 869-70 (2009).

lawyers, but as with Adam Smith's pin factory, a large firm allows lawyers in different positions to specialize in different tasks (whereas a solo practitioner or small firm lawyer must perform all tasks). The economist's assumption of profit maximization suggests the law firm will assign a task to the least expensive lawyer who can perform it at an acceptable level. Thus, assignment of tasks within the large firm provides insight on how law firms rank the complexity of tasks with the simplest tasks performed by the least experienced lawyers and the most complicated tasks performed by the most experienced ones.

Table 2
Distribution of Time on Tasks by Tenure in Tier One Firms

	Associates </= 2 Years	Associates >2 Years	All Partners	Tier One Total
Strong Employment Effects	**8.5%**	**4.5%**	**1.1%**	**4.1%**
Document Review	8.5%	4.5%	1.1%	4.1%
Moderate Employment Effects	**34.9%**	**38.5%**	**44.7%**	**39.0%**
Case Administration and Management	3.4%	2.4%	6.0%	3.7%
Document Drafting	4.4%	5.4%	4.8%	5.0%
Due Diligence	2.0%	1.6%	2.7%	2.0%
Legal Research	1.6%	0.4%	0.1%	0.5%
Legal Analysis and Strategy	23.5%	28.7%	31.1%	28.5%
Light Employment Effects	**56.3%**	**56.6%**	**54.2%**	**56.0%**
Document Management	0.7%	0.5%	0.2%	0.4%
Fact Investigation	13.9%	9.2%	6.7%	9.2%
Legal Writing	10.1%	12.5%	9.5%	11.4%
Advising Clients	8.3%	6.2%	14.8%	9.3%
Communications and Interactions	9.0%	11.1%	5.1%	8.8%
Court/Official Appearances and Preparation	12.0%	14.7%	13.8%	13.9%
Negotiation	2.4%	2.3%	4.2%	3.0%
*Totals***	99.7%	99.5%	100%	99.1%
Addendum: % of all Hours Billed by Tenure	18.0%	50.0%	32.0%	100.0%

** Percentages may not sum to 100% due to rounding.

1. The Relationship between Machine Complexity and Task Complexity

Table 2 reveals the absence of a strong association between the ease of automating a task (machine complexity) and whether the task is performed by a junior associate, a senior associate, or a partner (task complexity as viewed by the firm). Some data point toward a connection.

Document review is heavily computerized and when performed by firm lawyers (as opposed to contract attorneys), is largely performed by junior associates. Advising clients is difficult to computerize and much of it is performed by partners. If these were the only two data points, they would suggest that tasks with the lowest machine complexity are assigned to the least experienced lawyers, and tasks with the highest machine complexity are assigned to the most experienced lawyers. And this, in turn, would confirm the conventional wisdom that computers are having their greatest impact on the lowest level of lawyers within a firm. But the actual pattern is far less neat. The tasks of fact investigation and communication/interactions both have minimal computer penetration, and yet junior associates spend a greater percentage of their time on both tasks than do partners.

The factor that undermines a simple relationship between machine complexity and role within a firm is unstructured human interaction, a skill that has so far resisted automation but that is a part of lawyering tasks at every level.[96] The task of advising clients may require more experience than fact investigation, but both require an ability to conduct unstructured communication with other people—something junior associates and partners can do but computers cannot—which, in turn, illustrates that despite massive amounts of computing power, many tasks that are easy for humans are exceedingly difficult for computers.[97]

2. Putting Numbers on Employment Impacts

Estimating employment impacts requires translating "Strong," "Moderate," and "Light" employment effects into percentage reductions in lawyers' hours. Contrary to some popular

[96] Remus, *Reconstructing Professionalism, supra* note XX, at 33-34.
[97] This proposition explains why automation has historically had its greatest effect on "mid-skilled" jobs, such as assembly line and clerical work. It has had much less of an effect on the lowest wage jobs because those jobs involve both unstructured human interaction and unstructured physical movement. *See* Autor et al., *supra* note XX.

reports,[98] estimating computers' employment effects is a very imprecise process.[99] Acknowledging the imprecision, we nevertheless construct estimates by combining judgment based on interviews with two examples from among a limited set of studies on the effect of automation on other occupations.[100] These studies focus on computers' impacts on employee productivity (output per hour of labor), but if the volume of work is constant ("lump of labor"), a percentage gain in output per hour labor is equivalent to a percentage reduction in required hours of work.

Strong Employment Effects: Only one legal technology systematically falls within this category—automated document review, which was Markoff's original example. Automated document review continues to require senior lawyer time to train the software and review the results, and is not efficient for small classification problems. Nonetheless, to avoid underestimating automation's employment impacts, we assume that automated document review for discovery replaces 85 percent of all lawyer hours currently assigned to this task.

In theory, online dispute resolution could also fall into the category of heavy employment effects given that when used, it entirely replaces not only lawyers but judges as well. Its

[98] *See e.g.,* Carl Benedikt Frey & Michael A. Osborne, *The Future of Employment: How Susceptible are Jobs to Computerization?* (Sept. 17, 2013) (draft on file with the Oxford University Oxford Martin School).

[99] Among white-collar occupations, a cause of imprecision is the lack of good output measures that would allow measuring changes in employment holding output constant. Among blue-collar occupations, a cause of imprecision is the overlap between jobs that are being automated and jobs that are being sent offshore. Frank Levy & Richard J. Murnane, *How Computerized Work and Offshoring Shape Human Skill Demands,* in MARCELO SUAREZ-OROZCO ED., LEARNING IN THE GLOBAL ERA: INTERNATIONAL PERSPECTIVES ON GLOBALIZATION AND EDUCATION ch. 7 (Univ. of Cal. Press 2007).

[100] Autor, Levy, & Murnane, *supra* note XX, at 432; Sinan Aral, et al., *Information Technology and Information Worker Productivity: Task Level Evidence*, NBER Working Paper 13172 (2007); Susan Athey & Scott Stern, *The Impact of Information Technology on Emergency Health Care Outcomes*, 33 RAND J. Econ. 399 (2002); Ann Bartel, Casey Ichniowski & Kathryn Shaw, *How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills*, 122 QUART. J. ECON. 1721 (2007); J. Adler-Milstein & R. Huckman, *Electronic Health Record Use, Delegation, and Physician Productivity in Community Practices*, 19 Am. J. Managed Care 345-352 (2013).

potential to impact lawyer employment may be significant in the future,[101] but as discussed, its current use is limited. We therefore estimate that existing products, such as Modria, have a minimal impact on lawyer employment.[102]

Moderate Employment Effects: Moderate employment effects arise when a largely unstructured legal task has a significant structured component that can be computerized—for example, an automated precedent search, the structured part of due diligence or the question answering components of legal research. To calibrate the employment impact of this level of innovation, we refer to a case study of search-related innovation in exceptions processing at a large bank.[103] Exceptions processing requires determining the proper disposition of

> …checks written on accounts that have been closed, checks written for amounts greater than the balances in the accounts on which they are drawn, checks that customers request stop payments on, checks written for large amounts that require signature verification, and fraudulent checks.[104]

Each department employee reconciled a single type of exception. The work was made more complex because a single check could involve multiple exceptions. For example, individuals short of cash might buy time by writing multiple checks to creditors and by then submitting multiple stop-check orders. The result was substantial time spent both searching boxes of checks for particular items and coordinating work among employees addressing different exceptions for the same account.

When digital check images were substituted for paper checks in the workflow, exceptions employees gained rapid access to a particular check, resulting in reduced search time and, therefore, increased productivity. Simultaneously, the exceptions departments reorganized their

[101] *See supra* notes XX and accompanying text.
[102] The disputes that are resolved by these programs are generally small stakes ecommerce issues, for which it would not be economically feasible to hire a lawyer and litigate. *See supra* notes XX-XX and accompanying text.
[103] Footnote reference to Autor, Levy and Murnane, Upstairs, Downstairs (already cited in FN 19)
[104] *See* Autor, Levy & Murnane, *supra* note XX at 437.

workflow such that employees no longer focused on a particular type of exception but instead handled all exceptions for a particular set of accounts. This reorganization, which could have taken place with paper checks (though it did not), also increased productivity. The combined effect was to reduce the number of employees required to handle a constant volume of exceptions from 650 to 470—a reduction of 28 percent. A reasonable estimate attributes two-thirds of this reduction to the digitized images and one-third to the reorganization. Correspondingly, we assume that lawyering tasks in which computers have a Moderate Employment Effect reduce lawyer time devoted to those tasks by 19 percent.

Light Employment Effects: This category encompasses Fact Investigation, Legal Writing (as distinct from Legal Drafting), Advising Clients, Communications/Interactions, Court Appearances, and Negotiation.[105] These are tasks that entail largely unstructured work with limited room for automation.

To calibrate Light Employment Effects, we use a case study of a limited computer innovation in healthcare: Adler-Milstein and Huckman's study of the impact of electronic medical record (EMR) use on clinician productivity.[106] Productivity in the study is measured by "Relative Value Units" billed per clinician workday, which is the standard medical accounting measure of the volume and intensity of services provided. The study's sample consists of 42 medical practices, which were observed over three years during which they implemented EMR's at various rates. Findings indicate that one standard deviation in the use of an EMR increases clinician productivity by five percent. Services per patient visit did not increase, but physicians could see more patients per workday by using the EMR to delegate some services to physicians'

[105] Light employment effects also arise in tasks relating to document management. Because these tasks are usually performed by clerical staff, automation does not affect lawyer employment per se.
[106] Adler-Milstein & Huckman, *supra* note XX.

assistants. Relying on this example, we posit that adopting a computer innovation with Light Employment Effects would decrease required lawyer employment for a given task by five percent.

This leaves one set of technologies for which even the roughest estimation of employment impacts is exceedingly difficult——document templates sold directly to the public by firms like LegalZoom and Rocket Lawyer. Many commentators believe these templates will fully eliminate many lawyers' jobs.[107] There is reason to question such assertions, but it is difficult to achieve any level of certainty on either side of the issue. The critical question is the extent to which these services are tapping into a latent market of previously unserved individuals as opposed to taking business away from lawyers. LegalZoom representatives argue that it is overwhelmingly the former—that they serve individuals who would not otherwise have gone to a lawyer[108]—but it is of course in their interests to frame their business model as non-threatening to lawyers.

Indirect evidence of LegalZoom's impact on market share comes from its 2012 decision to table its planned Initial Public Offering after receiving insufficient interest from the markets.[109] Since that time, there is reason to think that LegalZoom's business has not grown as rapidly as it had projected.[110] This may be the result of regulatory responses from unauthorized

[107] *See supra* note XX.
[108] Telephone conversation with Eddie Hartman, Chief Product Officer, Legalzoom (September 18, 2015).
[109] Olivia Oran, *Update 1-LegalZoom IPO delayed*, REUTERS (Aug 2, 2012) available at http://www.reuters.com/article/2012/08/02/legalzoom-idUSL2E8J2EZF20120802. For Legalzoom's filed Form S-1, *see* http://www.sec.gov/Archives/edgar/data/1286139/000104746912005763/a2209299zs-1.htm.
[110] *See, e.g., The $425M LegalZoom deal is a win for VCs, but less exciting for the company or LA,* available at https://pando.com/2014/01/06/the-legalzoom-deal-is-a-win-for-vcs-but-less-exciting-for-the-company-or-la/ (last visited Oct. 28, 2015) (describing a $200 million investment by private equity firm Permira as "a bit 'meh,'" and the company's $425 million valuation in the deal as a "a slight downgrade from the $500 million-plus valuation…the company was most recently hoping to attract in the public markets [as part of the proposed IPO]"); *Has LegalZoom lost its bloom?*, available at http://www.theformtool.com/has-legalzoom-los-its-bloom/ (last visited Oct. 28, 2015) (describing that, prior to the proposed IPO, "LegalZoom was already experiencing the chill of a slowing growth rate and tighter margins in its traditional market, legal forms for sale").

practice of law committees, a topic that we address in Part III. But regardless of cause, the slowed growth offers reason to question sweeping conclusions about massive lawyer displacement. Because of all of these uncertainties, we include document templates marketed directly to the public under the general heading of document drafting—a task with moderate computer penetration, for which relevant technologies tend to replace parts but not all of a lawyer's job.

To summarize, our illustrative calculation rests on three estimates:

- Tasks where computer technology has a strong employment effect experience an 85 percent reduction in employment.
- Tasks where computer technology has a moderate employment effect experience a 19 percent reduction in employment.
- Tasks where computer technology has a light employment effect experience a 5 percent reduction in employment.

To calculate an overall employment impact, we first apply these percentages to the lawyers' use of time in 2014 in Tier 1 firms (Table 2).[111] By taking Table 2 as a baseline, we are assuming that: (i) all of these employment effects occur at one point in time, (ii) no employment effects have occurred previously, and (iii) there is no possibility of new work offsetting computer-driven reductions (the "lump of labor" assumption). If all technology were applied under this (admittedly unrealistic) scenario, we estimate lawyer employment would drop by slightly more than 13 percent. [112]

With the proviso that the 13 percent figure is a ballpark estimate, we can give the figure a more realistic (if less provocative) interpretation by assuming that the technology is phased in over five years. If there were no growth in the volume of legal work ("lump of labor") a 13

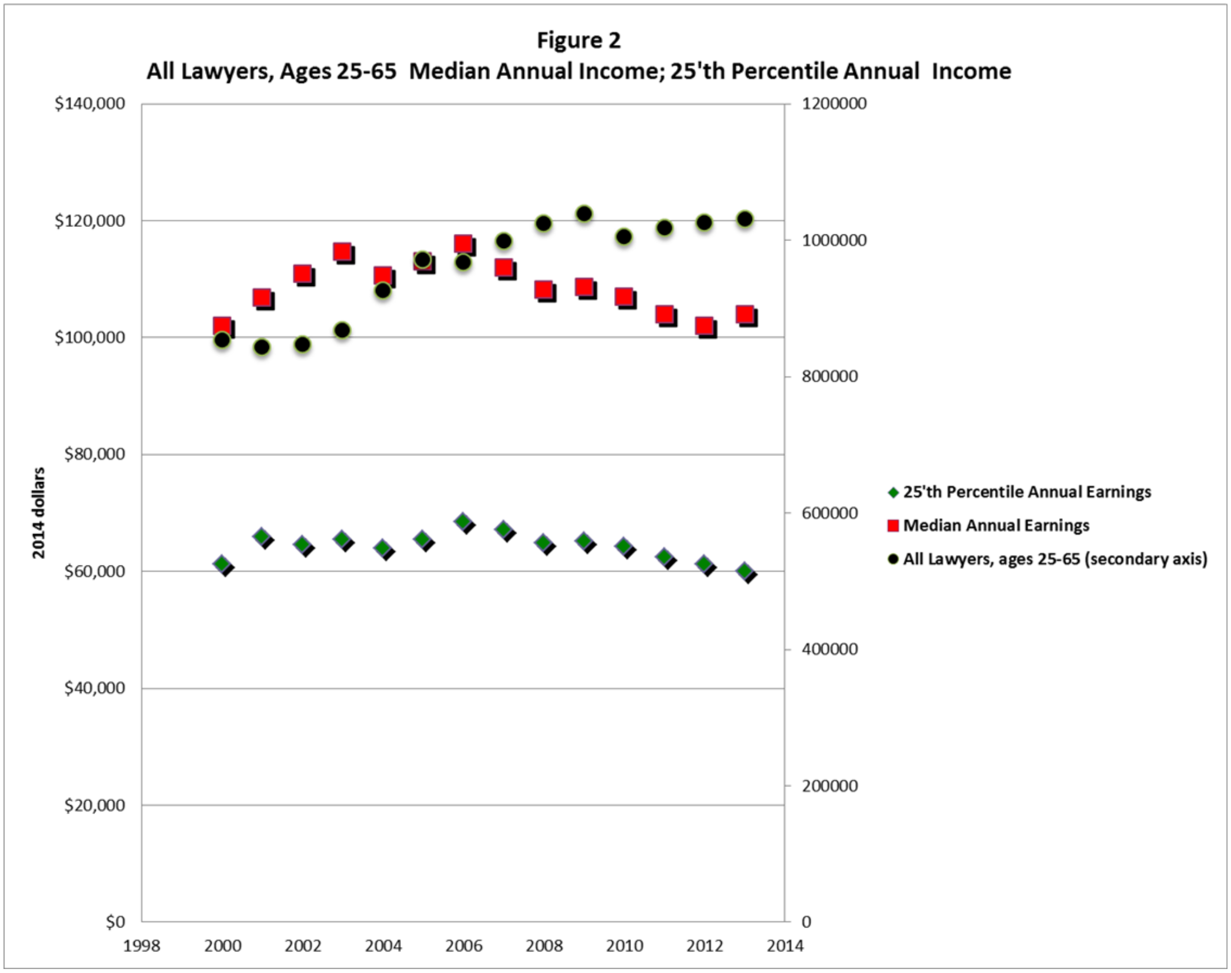
[111] Recall that the distribution of time on task in Tier2-Tier5 firms is similar to the distribution in Tier 1 firms.
[112] Our job loss estimate will be low if, prior to 2014, most document review work had been shifted to contract lawyers. In that case, automated document review would eliminate jobs of contract lawyers who are not included in the firm data of Table 2. Similarly, job loss our estimate will be low if solo practitioners spend much of their time on non-adversarial, formulaic issues that could be replaced by templates sold directly to individuals.

percent employment loss over five years would indicate lawyer productivity is increasing by 2.5 percent per year[113]. It follows that the volume of legal services must increase by at least 2.5 percent per year to offset automation's impact in reducing demand for lawyers' services. We note in passing that labor productivity increasing by 2.5 percent per year is a strong number: labor productivity for the entire U.S. non-farm business sector has grown at slightly less than 1.4 percent per year for the last ten years.

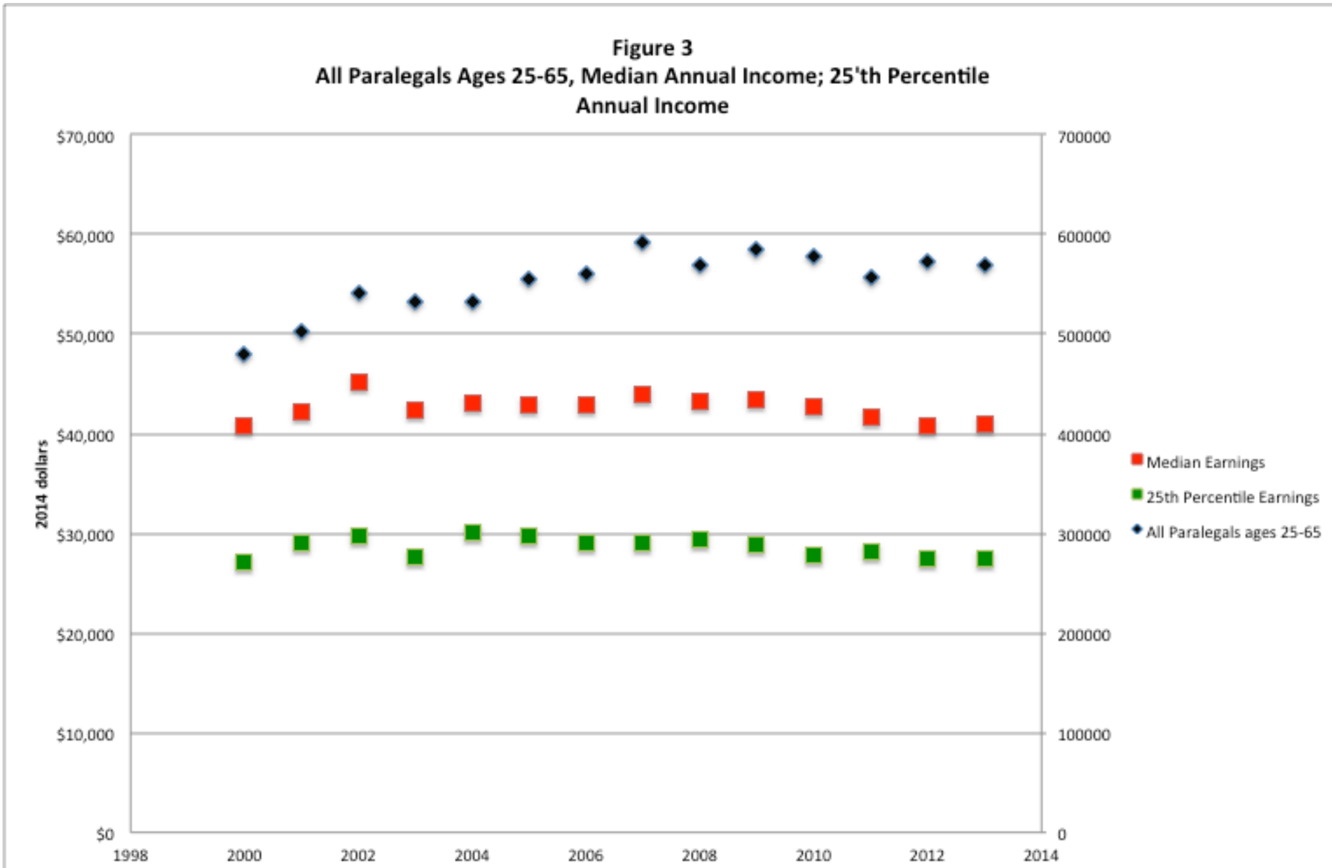
In reality there is no "lump of labor" and new work does continue to appear. The result is that automation acts as one more drag on a legal market that had already become relatively saturated by the middle of the last decade.

The condition of the market is illustrated in Figure 2, which displays median annual earnings for all lawyers ages 25-65 (men and women combined), and the total number of men and women lawyers, ages 25-65. Since roughly 2006, six years before the courts blessed predictive coding, and two years before the Great Recession, demand for lawyers could no longer accommodate continued increases in the supply of lawyers without steady reductions in median earnings.[114]

[113] We are indebted to Dan Sichel for this calculation.
[114] The data come from authors' tabulations of U.S. Bureau of the Census, *American Community Survey.* Figure 2 also includes 25'th percentile earnings to examine the possibility slowly declining median earnings obscures a collapse at the bottom of the earnings distribution. As shown in the figure, there is no evidence of this.



Compared to continued growth in the number of lawyers, the number of paralegals peaked in 2008 and declined slightly in subsequent years (Figure 3). This is consistent with the conjecture that document management and other productivity applications have had a greater impact on paralegal employment than on lawyer employment, but significantly more evidence would be required to prove the conjecture as fact.



* * *

Accordingly, our calculation shows that computers are affecting the market for legal services and will continue to do so as modeling abilities improve. Their impact, however, is not as drastic as headlines frequently suggest. Predictions of widespread labor displacement of lawyers by computers are overstated or, at the very least, premature. And yet, these predictions have been the nearly exclusive focus of those studying technology and the legal profession, distracting attention from the many other ways in which computers are impacting legal practice.

III. COMPUTERS, PROFESSIONALISM, AND THE RULE OF LAW

In this Part, we move beyond employment impacts and address distinct but equally important questions: How is technology changing (rather than replacing) the work of lawyers?

How should the profession address and regulate those changes? And what light do those changes shed on the value (or lack thereof) of organizing lawyers as a profession?

We start by briefly addressing a common refrain among critics of the profession—that the professional form is socially undesirable and economically inefficient. The profession, critics contend, is nothing more than institutionalized self-interest. We agree that the profession's current framework for addressing and regulating new technologies—unauthorized practice of law rules—appear motivated primarily by self-interest, but we contend that the answer cannot be to abandon all forms of regulation. Through a discussion of the ways in which computers approach various tasks differently than humans, and the wide-reaching ramifications of those differences, we show the importance of a comprehensive and nuanced normative and regulatory inquiry.

A. Skepticism of the Professional Form

Economists have long been skeptical of professional licensure. Milton Friedman observed half a century ago that "[l]icensure…frequently establishes essentially the medieval guild kind of regulation in which the state assigns power to members of the profession."[115] Such state-granted power, which takes the form of monopolistic market power and the right of self-regulation, allows licensees to advance their financial self-interest at the expense of the public interest.[116] Commenting on the medical profession, Friedman observed that licensure had

[115] MILTON FRIEDMAN, CAPITALISM AND FREEDOM (1962) ("Licensure…frequently establishes essentially the medieval guild kind of regulation in which the state assigns power to members of the profession.").
[116] *Id.* ("The most obvious social cost is that any one of these measures, whether it be registration, certification or licensure, almost inevitably becomes a tool in the hands of a special producer group to obtain a monopoly position at the expense of the rest of the public.").

"forced the public to pay more for less satisfactory medical service, and… retarded technological development both in medicine itself and in the organization of medical practice."[117]

Countless scholars and commentators, some with economic training and others from other scholarly traditions, have criticized the legal profession for bearing out this characterization of professionalism. They observe that throughout much of the twentieth century, the organized bar used barriers to entry to exclude women and minorities, and promulgated ethics rules that serve lawyers' self-interest at the expense of the public interest.[118] Particularly relevant to the topic at hand, these commentators observe that the profession's current framework for addressing new technologies—unauthorized practice of law statutes—have long been at the center of the bar's problematic behavior.[119]

These scholars' critique is powerful and persuasive, but it presents an incomplete picture of the legal profession. The organized bar has acted, and continues to act, in altruistic and public-serving ways at the same time that it acts in protectionist and self-serving ways.[120] Thus, if we abandon the professional form in its entirety in order to address its problems, we will

---

[117] *Id.*

[118] Seminal critiques of the profession along these lines came not from economists but from neo-Marxist sociologists. *See e.g.,* ANDREW ABBOTT, THE SYSTEM OF PROFESSIONS: AN ESSAY ON THE DIVISION OF EXPERT LABOR 184-86 (1988) (describing the American bar's efforts to solidify and expand its "jurisdiction" as against encroachment by non-lawyers); RICHARD ABEL, AMERICAN LAWYERS 40-48 (1989) (demonstrating the extent to which the profession's ethics rules protect lawyers rather than clients or the public); JERALD AUERBACH UNEQUAL JUSTICE: LAWYERS AND SOCIAL CHANGE IN MODERN AMERICA 88, 92, 99-102 (1976) (documented the organized bar's successful efforts, throughout much of the twentieth century, to use barriers to entry to exclude minorities, Jews, and women). More recently, a subsequent wave of scholars of the profession have reframed this critiques in economic terms. *See, e.g.,* RICHARD SUSSKIND, THE END OF LAWYERS? RETHINKING THE NATURE OF LEGAL SERVICES (2008); Gillian Hadfield, *The Cost of Law: Promoting Access to Justice Through the (Un)Corporate Practice of Law*, 38 Int'l Rev. L. & Econ. 43, 43 (2014); Benjamin H. Barton, *Why Do We Regulate Lawyers?: An Economic Analysis of the Justifications for Entry and Conduct Regulation*, 33 Ariz. St. L.J. 429, 457 (2001); Deborah Rhode, *Access to Justice: Connecting Principles to Practice*, 17 Geo. J. Legal Ethics 369, 371–72, 409 (2004).

[119] *See infra* notes ** and accompanying text.

[120] *See, e.g.,* Robert W. Gordon, *The Citizen Lawyer—A Brief Informal History of a Myth with Some Basis in Reality,* 50 Wm. & Mary L. Rev. 1169, 1183–84 (2009) (observing that throughout history, American lawyers have sometimes acted in publicly-oriented ways, and at other times in self-serving ways); KEITH MACDONALD, THE SOCIOLOGY OF THE PROFESSIONS 34-35 (1995) (observing that although some actions by the professions appear to be "mere self-enhancement," other actions consistently entail service to clients, patients, and society at large).

simultaneously lose significant and unique value that it creates for society.[121] We return to this point below, using automated legal services as a new window onto the core meaning and value of legal professionalism. Before we do, we address the more specific manifestation of the critique directed at UPL rules.

B. Existing Approaches

UPL rules, which limit the provision of legal services to individuals who are trained and licensed to practice law, are designed to protect the public from incompetent and unethical service providers.[122] These rules seek to distinguish tasks that can only be performed by trained and licensed lawyers from tasks that lay people, lacking the same training and ethical regulation, can nevertheless provide competently, reliably, and ethically.

For at least four reasons, this approach is not helpful. First, courts following this approach have posed for themselves an unanswerable question—is a given technology (generally an online service provider) more similar to a scrivener who completes a form by merely recording the information a customer relays (in which case the technology would not constitute UPL) or a service provider who aids in selecting and properly completing a form (in which case, it would be UPL)?[123] Neither alternative is ever clearly right or clearly wrong.[124] An online legal forms provider can be viewed as the functional equivalent of a mere scrivener insofar as it

[121] Among other things, the professional privileges uniquely facilitate the rule of law—without which, a free market cannot function. *See* Remus, Reconstructing Professionalism, at **.
[122] The principal justification prohibiting unauthorized practice of law is "to protect the public from the consequences of receiving legal services from unqualified persons." ABA MODEL RULES OF PROF'L CONDUCT [hereinafter M.R.] 5.5 annot., at 458 (ABA 2007) ("The proscriptions also facilitate regulation of the legal profession and protect the integrity of the judicial system.").
[123] *See, e.g.,* Janson et al. v. Legalzoom.com, Inc., 802 F.Supp. 2d 1053, 1059 (W.D. Mo. 2011) ("Plaintiffs urges the Court to follow the cases . . . which generally involve businesses providing a legal document preparation service for their customers . . . . Defendant Legalzoom argues that its website providing access to online document assembly software is the functional equivalence of [a] "do-it-yourself' divorce kit.").
[124] The court in *Janson* even acknowledged this, but declined to revise its analysis accordingly. *See id.* at ** ("None of the cases cited by the parties are directly on point, due to the novelty of the technology at issue here.").

is the user him or herself who enters the relevant information via the online questionnaire and completes the form,[125] or as the functional equivalent of a human service provider exercising judgment insofar as the software is programmed with deductive rules to ask the user a series of questions and, based on the answers, complete the appropriate document.[126]

Second, analogizing to human approaches also fails to appreciate that which is unique and different about legal technologies. Computers can be trained in ways that avoid human error such that we may be comfortable with a computer performing tasks we would not want performed by an untrained and potentially unreliable human.[127] And yet, reducing a lawyering task to a set of computer-implementable rules may over-simplify, ignore complexity, or create opportunities for error that are not immediately apparent.[128] We therefore may not want a computer performing particular tasks in all contexts, notwithstanding effective performance in one context.

A third problem stems from the poor fit between the UPL inquiry and technologies that lawyers use in representing clients (as opposed to those that are marketed directly to the public). Concluding that a non-lawyer cannot competently and reliably perform a particular task does not establish that a computer cannot help a lawyer do so. Perhaps for this reason, some commentators suggest that technologies that lawyers use and oversee are best addressed through

[125] *See id.* at 17 (noting LegalZoom's argument that "its customers—rather than LegalZoom itself—complete the standardized legal documents by entering their information via the online questionnaire to fill the document's blanks.").
[126] *See id. at* ** (observing that LegalZoom reassures consumers that "we'll prepare your legal documents," and that "LegalZoom takes over" once customers "answer a few simple online questions.").
[127] The *Janson* court ignored this, resting entirely on a formalistic UPL analysis. *Id.* at 20-21 ("Because those that provide [LegalZoom's] service are not authorized to practice law in Missouri, there is a clear risk of the public being served in legal matters by 'incompetent or unreliable persons.").
[128] *See infra* notes XX-XX and accompanying text.

the rules of lawyer oversight of non-lawyer service providers.[129] Applied to new technologies, these rules would permit adoption of new technologies where lawyers supervise their use and accept responsibility for their results. At least for now, however, few lawyers are sufficiently knowledgeable to oversee new legal technologies in a meaningful way.[130] Moreover, this approach suggests that computerizing all of a lawyer's functions would be permissible with oversight. But surely some tasks, such as in court advocacy and settlement or plea negotiations, cannot and should not be delegated to a computer.

A fourth and final problem with the UPL inquiry is that which we noted above—UPL prosecutions often appear to be self-interested efforts by the bar to protect its monopoly.[131] Scholars and commentators have long argued that non-lawyers can perform certain aspects of legal practice perfectly well, and that allowing them to do so would dramatically reduce the cost of legal services. Increasingly, commentators are extending the argument to computers,

---

[129] These rules, developed to address the outsourcing of work to non-lawyers or offshoring of work to foreign lawyers, provide that such activities are ethically permissible so long as the lawyer supervises the work and retains ultimate responsibility for the result. *See, e.g.*, ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 08-451 (2008) ("A lawyer may outsource legal or nonlegal support services provided the lawyer remains ultimately responsible for rendering competent legal services to the client under Model Rule 1.1."); *see also* Prof'l Ethics of the Fla. B., Op. 07-2 (2008) (approving of off-shore outsourcing); The Ass'n of the B. of the City of N.Y. Comm. on Prof'l and Jud. Ethics, Formal Op. 2006-3 (2006) (providing that a lawyer may outsource legal support services to overseas lawyers and non-lawyers if the lawyer supervises the work rigorously).

[130] A comment to Model Rule 1.1 advises lawyers of a professional duty to stay abreast of technological advances, *see* M.R. 1.1, cmt [5], but this provision has little teeth given the vagueness of its standard and its location in the comments rather than in an enforceable rule. Moreover, lawyers' generally low level of technical competency is reinforced by other provision of the Model Rules, which prescribe a reduced level of required oversight for automated legal work. *See, e.g.,* M.R. 5.3 cmt. [4].

[131] *See, e.g.,* DEBORAH RHODE, THE TROUBLE WITH LAWYERS, ch. 5 (Oxford U. Press 2015); Benjamin H. Barton, *Why Do We Regulate Lawyers?: An Economic Analysis of the Justifications for Entry and Conduct Regulation*, 33 ARIZ. ST. L.J. 429, 457 (2001); Roger C. Cramton, *Delivery of Legal Services to Ordinary Americans*, 44 CASE W. RES. L. REV. 531, 615 (1994); Deborah Rhode, *Access to Justice: Connecting Principles to Practice*, 17 GEO. J. LEGAL ETHICS 369, 371–72, 409 (2004); Deborah L. Rhode, *Professionalism in Perspective: Alternative Approaches to Nonlawyer Practice*, 22 N.Y.U. REV. L. & SOC. CHANGE 701, 709–13 (1996).

contending that if a computer can perform a task well, we no longer need lawyers to perform it.[132]

This reasoning was recently adopted by the U.S. Court of Appeals for the Second Circuit in a case construing the exemption from over-time pay for individuals "employed in a bona fide…professional capacity" under the Fair Labor Standards Act.[133] Plaintiff, a contract attorney, argued that document review, defined as "us[ing] criteria developed by others to simply sort documents into different categories," did not constitute the practice of law, such that he was not employed in a "professional capacity."[134] The Second Circuit agreed, reasoning that because these were "services that a machine could have provided," they could not possibly constitute the practice of law.[135]

Notwithstanding the reference to "services a machine could have provided," neither the Second Circuit, nor scholars and commentators expressing similar reasoning, engage with the difficult but critical inquiry of whether and how the machine approaches the task differently from a human. And yet, those differences have ramifications that extend beyond lowered costs, are central to a meaningful normative inquiry, and demonstrate the need for continued regulation.[136] As discussed, the computer's altered approach is often what makes automation attractive—it may sidestep opportunities for human error, improving accuracy and consistency. But it may

---

[132] *See, e.g.,* McGinnis & Pearce, *supra* note XX, at 3066; BARTON, *supra* note XX; Barton, *The Lawyer's Monopoly, supra* note XX, at 3068.
[133] Lola v. Skadden, Arps, Slate, Meagher & Flom LLP, No. 14-03845-cv, 2015 WL 4476828, at *2 (2d Cir. July 23, 2015) (reviewing an appeal from an order dismissing plaintiff's putative class action for violation of the Fair Labor Standards Act, 29 USC 201 et seq., for failing to pay overtime for document review).
[134] *Id.* at *6.
[135] *Id*. Plaintiff alleged that his work was closely supervised by the Defendants, and his "entire responsibility…consisted of (a) looking at documents to see what search terms, if any, appeared in the documents, (b) marking those documents into the categories predetermined by Defendants, and (c) at times drawing black boxes to redact portions of certain documents based on specific protocols that Defendants provided." *Id.* at *1 (internal quotations omitted).
[136] As noted above, *see supra* notes XX and accompanying text, it is far from established that all legal technologies will lead to lowered costs. *See also* Remus, *Predictive Coding, supra* note XX, at 1707.

also create new opportunities for error or have unintended consequences for legal practice. Access to deeply flawed and error-filled legal services cannot qualify as an acceptable, much less desirable, answer to the access to justice gap.

C. Moving Beyond UPL

UPL statutes offer an inadequate and problematic approach to evaluating and addressing new technologies, but the answer cannot be to abandon all forms of regulation. Here, we offer two illustrations of how important it is to identify and understand differences between a computer's and a human approach to a particular task—predictive coding and legal prediction software. These differences raise a host of issues not just for clients, but for the legal system more broadly, which professional regulation must address.

1. Predictive Coding

Predictive coding, the subject of the Second Circuit's decision, illustrates how a legal technology that eliminates error in some contexts may simultaneously create new risks or error in other contexts. To explain how and why, we begin with the different ways in which humans and computers approach the task of document review. A human lawyer examines a set of documents page-by-page to identify relevant meaning and content. Predictive coding, in contrast, identifies particular combinations of document features pursuant to statistical probabilities of relevance, with no reference to meaning. As described above, this is not a problem when the goal is to locate types of data and information that have been well specified in advance, such as in discovery practice.[137] But it makes the software far less effective when the goal is to identify unusual or unexpected information, such as in the unstructured aspects of the due diligence that

---

[137] *See, e.g.,* Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 RICH. J.L. & TECH. 1, 3 (2011); Herbert L. Roitblat et al., *Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review*, 61 J. AM. SOC'Y FOR INFO. SCI. & TECH. 70, 70, 74–75 (2010).

precedes a corporate transaction.[138] As noted above, some firms are making progress in automating aspects of due diligence but at least for the time being, their products are effective only in reviewing large volumes of similar documents.

Moreover, even within discovery practice, predictive coding may create new risks of error by failing to recognize "hot documents"—documents that are highly relevant and damaging to the producing party. Such documents, which generally prove, explain, or describe significant decisions or events related to the litigation, frequently employ unusual language, syntax, or even coded language (because individuals often change their normal writing styles, becoming particularly formalistic or vague, when they explain major decisions or acquire potential liability). As a result, the most relevant documents in a case may use language and tone that the software, trained on a sample of normal documents, will fail to recognize.[139] The problem is virtually intractable if the author reverts to coded language.[140] And yet, because the machine-learning algorithm does not recognize the existence of a problem in these situations—because it has encountered but simply ignored an unanticipated contingency—it will not give the user a warning.

Predictive coding software may also increase the harm that flows from any single instance of human error. Such error, which would be made in coding the initial sample set of documents and training the computer, would then be projected across the entire document set. If and when problematic results are identified late in the process, it will likely be impossible to differentiate between human and computer error. So long as document production is considered the practice of law and/or lawyers retain supervisory responsibility over its outsourcing, lawyers

---

[138] *See supra* notes XX-XX and accompanying text.
[139] Telephone conversation with Nathalie Hofman, Huron Consulting (July 21, 2015).
[140] For example: "I bought two pounds of flounder at the fish market today" to indicate a completed transaction. *Id.*

will remain accountable for any problems that arise, regardless of source.[141] But if document review is carved off from the practice of law and lawyers are no longer held responsible for work outsourced to technology vendors, it is unclear how accountability will be allocated between lawyers and software providers—or whether there will be meaningful accountability at all.

Our point is not that predictive coding has no value or should never be used, but rather, that its apparent efficacy in some contexts does not alone support its use in all contexts. Nor does it support carving document review out from the purview of the profession. Predictive coding increases the risk for error some cases, at the same time that it decreases the risk in others. It can therefore be a useful complement to legal practice, but only with sufficient understanding and oversight.

2. Legal prediction software

A second example, legal prediction software, illustrates ramifications that extend beyond an individual case or client. Again, our starting point is the ways in which human and computer approaches differ—here, to the task of informing and guiding a client's decision-making processes. Lawyers generally meet with the client to elicit relevant information about the client's situations and goals.[142] They gather relevant information by, among other things, examining relevant opinions and other sources of law, seeking advice and information from colleagues, and investigating judges' and other law-makers' sympathies and tendencies.[143] They then filter the resulting information through their legal training, their understanding of the

---

[141] *See supra* note XX, and accompanying text.

[142] DAVID A. BINDER, ET AL., LAWYERS AS COUNSELORS: A CLIENT-CENTERED APPROACH 2-15 (1991); Kruse, *Beyond Cardboard Clients, supra* note XX, at 104; Kruse, *Promise of Client-Centered Professional Norms, supra* note XX, at 346; Binder (2012).

[143] *See,* Theodore W. Ruger et al., *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150, 1183 (2004) (noting that experts retained to make Supreme Court predictions reported relying on "traditional legal materials, such as court decisions, statutes, and the briefs in the case," and "[s]omewhat less frequently . . . reported that they had read scholarly commentary or spoken with colleagues before reaching their prediction.").

client's interests, and their professional commitments, and give advice that advances their clients' interests, considered holistically, consistent with the constraints of the law. As described above, software prediction programs take a Big Data approach, collecting and analyzing massive amounts of data about judges and their decisions (far more than a human lawyer could ever digest).[144] The output, like that of predictive coding, is expressed in the form of statistical probabilities.

Existing evidence suggests that legal prediction software has higher levels of accuracy than human prediction[145] (and more broadly, that statistical prediction is more accurate than clinical prediction in most contexts[146]). But if such software completely displaces lawyers, the increased accuracy would be accompanied by a number of detrimental consequences. For one thing, reducing advice to prediction would eliminate a core function of lawyering—counseling compliance with the law. If a client's only legal advice comes from a computer's prediction of how a court will likely respond, advising will become calculating what a client can get away with, instantiating the Holmesian Bad Man view of the lawyer.[147]

A useful example comes from tax practice, where lawyers' traditional refusal to design or endorse arrangements that lack any economic substance has been critical to the integrity and

---

[144] *See supra* notes XX-XX and accompanying text.
[145] *See* Ruger et al., *supra* note XX (reporting that a statistical model, which relied on general case characteristics predicted 75 percent of the Court's affirm/reverse results correctly, while legal experts collectively got 59.1 percent right); *see also* Elizabeth Earl*, Law profs develop Supreme Court predictor to better understand court decisions,* ABA JOURNAL (Dec. 01, 2014) available at http://www.abajournal.com/magazine/article/law_profs_develop_supreme_court_predictor/?utm_source=maestro&utm_medium=email&utm_campaign=tech_monthly (last accessed Oct. 25, 2015).
[146] *See* William M. Grove, *Clinical versus Statistical Prediction: The Contribution of Paul E. Meehl*, 61 J. CLIN. PSYCH. 1233 (2005).
[147] *See* Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457 (1897); Robert Cooter, *The Legal Construction of Norms: Do Good Laws Make Good Citizens? An Economic Analysis of Internalized Norms,* 86 VA. L. REV. 1577, 1591 (2000) ("[T]he 'bad man' treats the law as 'external,' to himself, in the sense that he considers it to lie outside of his own values. Economic models of law typically accept the 'bad man' approach and add a rationality element to it: a rational 'bad man' decides whether or not to obey the law by calculating his own benefits and costs, including the risk of punishment.").

proper functioning of our tax system.[148] As part of this, the elite tax bar historically eschewed tax shelter practice as unethical.[149] But consensus broke down through the 1990s as the increasingly fluid market for legal services led to an overwhelming focus on profits per partner.[150] Individual tax lawyers began designing and marketing abusive arrangements that ultimately cost the American public hundreds of millions of dollars in tax revenues.[151] The incident is widely regarded as a massive breakdown in the proper role of the lawyer and the tax bar.[152]

Computerized tax advice, divorced from a lawyer's ethical obligations, could easily create a new tax shelter scandal. By accurately predicting the chances of auditing and detection, the computer would give a rational taxpayer necessary information to determine when a particular arrangement's promised tax savings would outweigh the probability of detection combined with the penalty if detected. By eliminating the lawyer's role as both a gatekeeper (refusing to facilitate such transactions) and advisor of compliance (highlighting additional costs of noncompliance, such as damage to the taxpayer's reputation if underpayment is detected and potential guilt from evasion), such software would eliminate necessary checks on such an approach to tax obligations.[153]

---

[148] *See, e.g.,* TANINA ROSTAIN & MILTON C. REGAN, JR., CONFIDENCE GAMES: LAWYERS, ACCOUNTANTS, AND THE TAX SHELTER INDUSTRY 73 (MIT Press 2014).
[149] *Id.; see also* Wendel, *Professionalism as Interpretation*, *supra* note XX, at 1171-72 (2005).
[150] ROSTAIN & REGAN*, supra* note XX, at 73; 217–18; *see also* Peter C. Canellos, *A Tax Practitioner's Perspective on Substance, Form and Business Purpose in Structuring Business Transactions and in Tax Shelters,* 54 SMU L. REV. 47, 51-52 (2001) (identifying two "vastly different" communities of practitioners: the tax bar and the tax shelter bar.)
[151] *See generally* ROSTAIN & REGAN*, supra* note XX.
[152] *Id.;* Wendel, *Professionalism as Interpretation*, *supra* note XX.
[153] *See id.* at 1172 (If lawyers were permitted to adopt a Holmesian bad man interpretive attitude…then the law would end up having no boundaries at all...This conclusion may sound unduly apocalyptic, but experience with practices such as tax sheltering shows that lawyers can always plan around formal legal norms, given sufficient resources and creativity."); Yoram Keinan, *Playing the Audit Lottery: The Role of Penalties in the U.S. Tax Law in the Aftermath of* Long Term Capital Holdings v. United States, 3 BERKELEY BUS. L.J. 382 (2006).

More broadly, reducing legal advising to legal prediction would also threaten to impede the law's development. Predictability and stability are of course critical rule of law values, but so too is democratic participation in law-making.[154] A core way in which citizens participate is through their lawyers, who translate their interests into persuasive and sometimes novel arguments as to how the law should apply to their clients' circumstances. Lawyers can do so because our legal system is about reasons as well as outcomes—reasons, asserted by lawyers and memorialized in judicial opinions, which provide a continual opportunity through which to debate and potentially change the law.[155] But if lawyering is replaced by computer prediction, we will shift to a system that is more about outcomes than reasons—and outcomes that are inescapably "informed by the world as it was in the past, or, at best, as it currently is."[156]

Of course, this may change over time. As natural language processing capabilities advance and computers become more capable of processing concepts and analogies, computer prediction may be joined by computer creativity, in the form of combinatorial processing. As lawyers recognize, creativity and novelty in legal argument generally comes from importing legal concepts from one area of law into another, and by combining existing arguments in new and persuasive ways. Indeed, knowledge production in many fields proceeds in this way—by

---

[154] Jeremy Waldron, *The Concept and the Rule of Law*, 43 GA L REV 1, 5 (2008) ("[O]ur understanding of the Rule of Law should emphasize not only the value of settled, determinate rules and the predictability that such rules make possible, but also the importance of the procedural and argumentative aspects of legal practice."); *see also* Benjamin Ewing & Douglas A. Kysar, *Prods and Pleas: Limited Government in an Era of Unlimited Harm*, 121 YALE L.J. 350 (2011) ("[T]here is another side to the value of the rule of law that is especially significant in the adversarial American system: law as a structured discourse in which individuals are entitled to articulate their grievances or face their accusers, to stake their claims, and to advance reasons in support of them.").

[155] Frederick Schauer, *Giving Reasons*, 47 STAN. L. REV. 633, 658 (1995) ("That giving reasons is a way of opening a conversation may in fact be an independent basis for a reason-giving requirement."); Ruger et al., *supra* note xx, at 1193 (noting that the Supreme Court's "role in American society is not merely to process important disputes expeditiously. Rather, the ways in which it addresses those disputes—not merely through outcomes, but through its rationales, its analytical framework, and its language—both gives voice to certain values and influences public understanding of these issues.").

[156] C. Anderson, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, WIRED 16.07 (2012); Martin Hilbert, *Big Data for Development: From Information- to Knowledge Societies*, SSRN Scholarly Paper No. ID 2205145 (2013), available at http://papers.ssrn.com/abstract=2205145.

recombining existing ideas in new and innovative ways.[157] Computers cannot currently do this, but their ability to do so will likely increase over time. Much as medical diagnostic programs currently suggest disease hypotheses to physicians based on patient symptoms, legal argument programs may be able to suggest new and promising combinations of existing arguments tailored to a client's factual circumstance. For now, however, computer programs are highly effective in making predictions given the legal system as it currently exists, and far less so in making suggestions for how the legal system could or should evolve.

Much of the prediction process is hidden from view. Like Big Data applications generally, most programs give a user results without showing the precise combination of factors that produced those results.[158] Certainly, an application's programmers can view the code of the relevant inductive rule model, but the code is not always interpretable by the programmer, much less a lay person, and will frequently be proprietary, protected as a trade secret. Interpretability could be made a priority, such that no outcome would be accepted—whether by a client, a lawyer, or a court—without a full explanation of inputs. But there are at least two reasons to doubt that this will happen. First, requiring every outcome to be accompanied by a complete explanation of inputs (features that gave rise to the computer's model) would be exceedingly expensive and time consuming. Most users would not be willing to bear that expense. Second, demanding interpretability would likely hold the technology back. Interpretability might be a reasonable goal for applications that consider a modest amount of data, but as the universe of

[157] *See* Martin L. Weitzman, *Recombinant growth*, 113 Q. J. of Econ. 331, 331 (1998) ("Production of new ideas is made a function of newly reconfigured old ideas in the spirit of the way an agricultural research station develops improved plant varieties by cross-pollinating existing plant varieties").

[158] David Martens & Foster Provost, *Explaining Documents' Classification* 2 (N.Y.U. Stern Sch. of Bus., Working Paper No. CeDER-11-01, 2011), http://archive.nyu.edu/handle/2451/29918 ("Unfortunately, due to the high dimensionality, understanding the decisions made by the document classifiers is very difficult. Previous approaches to gain insight into black-box models do not deal well with high-dimensional data."). *See also* Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1520 (2013) ("Yet interpretability has a flip side as well. Mandating interpretability might render the process less complex and therefore less accurate.").

data expands to tens of thousands or millions of variables (words, linguistic features, data points), the goal of interpretability becomes more and more difficult, if not impossible.

This lack of transparency threatens worrisome consequences over time. If clients increasingly rely on software predictions in determining a course of action—in deciding, for example, whether to file a complaint, to defend a case, or to pursue a particular corporate transaction—the software predictions, by virtue of their influence over conduct, will influence the law in action. Without anyone realizing it, factors encoded into those predictions—including discriminatory or otherwise problematic factors—could then become encoded into broader swaths of law. For example, a computer might discover a weak correlation between a particular court's decisions and the gender and ethnicity of the litigants. The estimated statistical model would then account for the correlation in predicting success or failure. Because the correlation is weak, the model's results might not immediately alert us to its influence in a way that would allow for accountability. Nevertheless, the discriminatory pattern would inform predictions of the court's decisions, and litigant behavior in the shadow of those decisions.

Finally, there is value in the explanations that lawyers give to their clients about why they are proposing the course of action they are proposing, just as there is value in judges explaining the results they reach. In discussing, explaining, and taking responsibility for those reasons, a lawyer may increase the client's understanding and awareness of the way the legal system functions and the client's place within it. The lawyer may also give the client comfort and confidence in the proposed course of action, and enhance the client's perceptions of fairness and satisfaction with the system.[159] This may not be so for all clients and all litigants. Some may not

[159] E. ALLAN LIND & TOM R. TYLER, THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE 64–65, 219-20 (1988).

want or try to understand what their lawyers tell them, but many clients do and will. They will lose this opportunity if all they are given is a computerized prediction.

D. The Core of Professionalism

The foregoing discussion shows that new legal technologies cannot be addressed and regulated through UPL statutes, but nor can they be accepted without any evaluative and regulatory inquiry. Computers approach tasks differently than humans and those differences must be carefully examined in determining how new technologies can and should be used, and whether new technologies can help to address the access to justice gap.

In this final section, we shift from this point to a broader one—that carefully considering new technologies sheds light not only on regulation of particular technologies, but on the value and significance of professional regulation generally. To show how and why this is so, we offer a simple thought experiment. Suppose that new software can accurately predict the likelihood that an individual will be audited by the Internal Revenue Service (IRS) and, if audited, that the proposed tax treatment of an asset-sheltering trust will be upheld. The software offers each prediction as a numerical probability, and there are no error costs. It is marketed to, and widely adopted by, financial planners who serve wealthy clients interested in minimizing gift and estate taxes.

What will be lost if this software eclipses the advice of tax and estate planning lawyers? Answering this question highlights value that lawyers provide and that, at least for the time being, computers cannot.

- <u>Counseling</u>. The tax software can predict how the IRS will act, but it cannot and will not counsel the taxpayer on how to proceed, including on the value of compliance and the possibility of an alternative course of action.[160] Nor can it push back against a taxpayer

[160] Deborah L. Rhode, *The Profession and the Public Interest*, 54 STAN. L. REV. 1501 (2002) ("One of the most crucial functions of legal counsel is to help individuals evaluate short-term economic objectives in light of long-term

who insists on proceeding with an illegal scheme, notwithstanding the fact that, as Elihu Root famously asserted, sometimes the proper role of the lawyer is to tell clients "that they are damned fools and should stop."[161]

- A robust understanding of law. Individuals planning their affairs pursuant to the software's prediction will come to experience the law purely in terms of what conduct will and will not be sanctioned—"what the courts will do in fact, and nothing more."[162] This impoverished view of the law will have detrimental consequences not only for compliance,[163] but for perceptions of the legal system's legitimacy and democratic participation in law-making.[164]

- Respect for clients' interests. The software objectifies a user[165] by assuming that the objective of all users is to use any asset-sheltering trust arrangement for which the projected savings outweigh the risk of detection. Some individuals engaged in estate planning seek excessively aggressive strategies, but others simply want to ensure that they are not needlessly sacrificing assets that could be shielded under well-settled law. The software simply ignores this, projecting one set of interests onto all clients.[166]

---

reputational concerns and to live up to their best, not worst instincts."); Robert W. Gordon, *The Independence of Lawyers*, 68 B.U. L. REV. 1, 18 (1988); Harold Williams, *Professionalism and the Corporate Bar*, 36 BUS. LAW. 165-66 (1980).

[161] Deborah L. Rhode, *Moral Counseling*, 75 FORDHAM L. REV. 1317, 1321 (2006) (quoting Philip C. Jessup, 1 Elihu Root 133 (1938)); *see also* DEBORAH L. RHODE, IN THE INTERESTS OF JUSTICE (2000); Wendel, *Professionalism as Interpretation, supra* note XX; Robert W. Gordon, *A New Role for Lawyers?: The Corporate Counselor After Enron*, 35 CONN. L. REV. 1185 (2003); Tanina Rostain*, Ethics Lost: Limitations of Current Approaches to Lawyer Regulation,* 71 S. CAL. L. REV. 1273, 1274-75 (1998).

[162] Holmes, *supra* note XX, at 460-61.

[163] Robert W. Gordon, *A Collective Failure of Nerve: The Bar's Response to Kaye Scholer*, 23 LAW & SOC. INQUIRY 315 (1998) ("[T]he order of rules and norms, policies and procedures, and institutional actors and roles that make up the legal system . . . is only as effective as voluntary compliance can make it; for if people routinely start running red lights when they think no cop is watching (or hire lawyers to keep a lookout for the cops, and to exhaust the resources of traffic courts arguing the lights were green), the regime will collapse."); Stephen Pepper, *Counseling at the Limits of the Law: an Exercise in the Jurisprudence and Ethics of Lawyering,* 104 YALE L. J. 1545, 1547-48 (1995) ("In a complex legal environment much law cannot be known and acted upon, cannot function as law, without lawyers to make it accessible to those for whom it is relevant."); W. Bradley Wendel, *Legal Ethics As "Political Moralism" or the Morality of Politics*, 93 CORNELL L. REV. 1413, 1417-18 (2008)

[164] Wendel, *Professionalism as Interpretation, supra* note XX, at 1167 ("[T] he law cannot operate as a device to settle normative conflict and coordinate activity without a commitment on the part of law-interpreters to respect the substantive meaning standing behind the formal expression of legal norms.")

[165] Simon, *supra* note XX, at 53-54 (warning that lawyers who adhere to the dominant ideology of professionalism "impute certain basic aims to the client," which tend to be legalistic and to "emphasize extreme selfishness."); Kruse, *The Promise of Client-Centered Professional Norms, supra* note XX, at 346 ("This tendency to "issue-spot" clients and construct their objectives solely in terms of maximizing legal interests can cause lawyers to minimize the importance of the other cares, concerns, commitments, relationships, reputations, and values with which the clients legal interests are intertwined.").

[166] And yet, as Kate Kruse and others have persuasively argued, lawyers can and should work to advance and represent their clients' interests, understood holistically; not the interests that they or the legal system project onto clients Kruse, *Beyond Cardboard Clients*, *supra* note XX, at 127-28 (describing client-centered lawyering, which entails "hearing clients' stories and understanding their values, cares, and commitments," as an answer to the problem of legal objectification);. BINDER, ET AL., *supra* note XX, at 2-15.

- Access to reasons. The tax software, like most Big Data applications, offers no reasons for its predictions. And yet, reasons, from lawyers as much as from judges,[167] are a critical source of both stability and change in the law[168] and a critical expression of respect for participants in the legal system.[169] Without reasons, neither the taxpayer nor the financial planner could understand the law so as to follow it or extrapolate the result to similar cases.[170] Nor could they critique the result, or argue for change.[171]

- Interaction with the legal system. Finally, widespread displacement of estate and tax lawyers by prediction software would eliminate a critical mechanism through which the state and society interact.[172] Lawyers translate their clients' interests into terms the legal system can understand and act upon, and the law into terms that their clients can understand and act upon.[173] Here, a lawyer could educate a taxpayer regarding the IRS's regulatory goals, and suggest an arrangement that would still minimize taxes without thwarting those goals. Or the lawyer could represent the taxpayer's interests in challenging the IRS's treatment of a particular arrangement or interpretation of a particular Code provision.

These features of lawyering, which we would lose in the automation of legal services, are deeply rooted in, and essential for, the rule of law. The rule of law necessitates respect for and compliance with law from a variety of sources even absent active enforcement.[174] It entails respect for the autonomy and dignity of citizens, including their self-defined interests.[175] It requires reasons, both as a resource for stability and a mechanism for change,[176] and it requires

---

[167] Wendel, *Interpretation as Professionalism, supra* note XX, at 1169-70 (discussing professionalism as "demand[ing] that lawyers provide a public, reasoned justification for an interpretation of legal texts one which is plausible in light of the interpretive understandings of a professional community.").
[168] David Luban, *Natural Law as Professional Ethics: A Reading of Fuller*, in NATURAL LAW AND MODERN MORAL PHILOSOPHY 176, 204 (Ellen Frankel Paul, Fred D. Miller, Jr. & Jeffrey Paul eds., 2001).
[169] *Id.* at 656; Luban, *Natural Law as Professional Ethics, supra* note XX, at 110-11 (discussing Fuller's distinction between law and managerial direction, and view that the former implies "a certain built-in respect for [the] human dignity" of those subject to the law).
[170] Schauer, *Giving Reasons*, *supra* note XX, at 641 ("When we provide a reason for a particular decision, we typically provide a rule, principle, standard, norm, or maxim broader than the decision itself, and this is so even if the form of articulation is not exactly what we normally think of as a principle.").
[171] *Id.* at 658.
[172] *See* Daniel Markovits, A Modern Legal Ethics: Adversary Advocacy in a Democratic Age 171-200 (Princeton U. Press 2008).
[173] *Reconstructing Professionalism*, *supra* note XX, at 37.
[174] Gordon, *A Collective Failure of Nerve*, *supra* note XX, at 315; Robert W. Gordon, *The Role of Lawyers in Producing the Rule of Law: Some Critical Reflections*, 11 THEORETICAL INQUIRIES L. 441, 442 (2010).
[175] Luban, *supra* note XX.
[176] *Id.*

participation in the development and application of law, and its evolution over time.[177] These are things that lawyers uniquely ensure and support, but that computers cannot and do not.

To be sure, these costs exist in significant tension with potential benefits, also central to the rule of law, including increased certainty through greater determinacy in the law and increased access through lowered prices. Moreover, they would be costs of eliminating lawyers entirely, and not a necessary consequence of the technologies themselves. The import of our thought experiment is not, therefore, to condemn all legal technologies. Rather, it is to show that the normative inquiry regarding new technologies and the associated regulatory inquiry regarding how we should manage their use must be at once detailed (rooted in the technology) and far-reaching (accounting for systemic effects). It cannot be shoe-horned into the UPL framework and answered with a surface level comparison of a computer's performance to human performance. Nor can it be sidestepped by reference to the importance of lowering costs and increasing access.

Our thought experiment has another benefit. Through a new lens and from a new perspective, it sheds light on the meaning of legal professionalism. Those things that computers cannot provide, notwithstanding massive computing power—counseling, a robust understanding of law, an appreciation of client interests, the provision of reasons, and interaction with the legal system—are core professional values. They are also central to the rule of law, illustrating a fundamental and necessary connection between the two.[178]

---

177 Waldron, *supra* note XX, at 5; Ewing & Kysar, *supra* note XX, at 350; Remus, *Reconstructing Professionalism, supra* note XX, at 34-38.

178 David Luban, *Natural Law as Professional Ethics, supra* note XX, at 100 ("the rule of law relies on the professional ethics of lawyers."); TERENCE C. HALLIDAY, BEYOND MONOPOLY: LAWYERS, STATE CRISES, AND PROFESSIONAL EMPOWERMENT 370–71 (1987) (discussing ways in which legal professions take primary responsibility for sustaining and advocating the integrity of the legal process and the rule of law throughout the world).

CONCLUSION

We have argued that most writing on the computerization of legal services overstates the likely employment impacts. Certainly, automation is having a significant impact on the labor market for lawyers and that impact will increase over time, but predictions of imminent and widespread displacement of lawyers are premature. A careful look at existing and emerging technologies reveals that it is only relatively structured and repetitive tasks that can currently be automated. These tasks represent a relatively modest percentage of lawyers' billable hours.

We have also argued that the existing literature focuses too narrowly on employment impacts, ignoring an important set of broader questions. The broader inquiry starts with the ways in which computers approach particular tasks differently than humans, and then asks how those differences may change legal practice and through it, the law itself. These questions are critical to a meaningful normative and regulatory approach to new technologies, and will only become more pressing as legal technologies continue to advance.