We Robot Readers—
We look forward to your comments and suggestions on this partial draft.

AI and professional work: The practice of law with automated decision support technologies

Daniel N. Kluttz
Deirdre K. Mulligan*
University of California, Berkeley School of Information

*Corresponding author:
Deirdre K. Mulligan
School of Information
University of California, Berkeley
303B South Hall
Berkeley, CA 94720-4600
dmulligan@berkeley.edu

***Authors listed alphabetically by last name*.**

We Robot 2019
April 12-13, 2019
University of Miami School of Law
Coral Gables, FL

ABSTRACT

Technical systems employing algorithms are shaping and displacing human decision making in a variety of fields. As technology reconfigures work practices, researchers have documented potential loss of human agency and skill, confusion about responsibility, diminished accountability, and both over- and under-reliance on decision-support systems. The introduction of predictive algorithm systems into professional decision making compounds both general concerns with bureaucratic inscrutability and opaque technical systems as well as specific concerns about encroachments on expert knowledge and (mis-)alignment with professional logics, liability frameworks, and ethics. To date, however, we have little empirical data regarding how automated decision-support tools are being debated, deployed, used, and governed in professional practice.

The objective of our ongoing empirical study is to analyze the organizational structures, professional rules and norms, and technical system properties that shape professionals' understanding and engagement with such systems in practice. As a case study, we examine decision-support systems marketed to legal professionals, focusing primarily on technologies marketed for e-discovery purposes. Generally referred to as "technology-assisted review" (TAR), and more specifically as "predictive coding," these systems increasingly rely on machine-learning techniques to classify and predict which of the voluminous electronic documents subject to litigation should be withheld or produced to the opposing side. We are accomplishing our objective through in-depth, semi-structured interviews of experts in this space: the technology company representatives who develop and sell such systems to law firms and the legal professionals who decide whether and how to use them in practice. We report research insights about how these systems and the companies offering them are shaping relationships between lawyers and clients, how lawyers are grappling with professional obligations in light of these shifting relationships, and the ways these systems construct and display knowledge. We argue that governance approaches should seek to put lawyers and decision-support systems in deeper conversation, not position lawyers as relatively passive recipients of system wisdom who must rely on out-of-system legal mechanisms to understand or challenge them. This requires attention to both the information demands of legal professionals and the processes of interaction that elicit human expertise and allow humans to obtain information about machine decision making.

INTRODUCTION

As applications based on advancements in fields such as cloud computing and machine learning have spread to the workplace, scholars and legal commentators have debated the extent to which such technical systems will affect markets for legal services, the practice of law, and the legal profession.[1] AI-based systems aimed at automating or assisting in lawyerly tasks and decision making are currently being employed in a wide range of practice domains—contract drafting and review, due diligence in mergers & acquisitions, risk-assessment in criminal justice settings, legal search and research, and document analysis and review in e-discovery, to name a few.[2] Technology-assisted review (TAR), also called "predictive coding," systems for the discovery phase of litigation provide an interesting example of a machine learning based decision support system infiltrating a professional domain.

Our research explores how professional identity, interactions with clients and vendors, and organizational structures are shaping the adoption, use, and perceptions of TAR systems in the field of

---

[1] See, e.g., John Markoff, *Armies of Expensive Lawyers, Replaced by Cheaper Software*, THE NEW YORK TIMES, Mar. 4, 2011, https://www.nytimes.com/2011/03/05/science/05legal.html; RICHARD SUSSKIND & DANIEL SUSSKIND, THE FUTURE OF THE PROFESSIONS: HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS (2015); Daniel Martin Katz, *Quantitative Legal Prediction - Or - How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909 (2013); Dana A. Remus & Frank Levy, *Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law*, 30 GEO. J. LEGAL ETHICS 501 (2017); Tanina Rostain, *Robots versus Lawyers: A User-Centered Approach*, 30 GEO. J. LEGAL ETHICS 559 (2017); Sean Semmler & Zeeve Rose, *Artificial Intelligence: Application Today and Implications Tomorrow*, 16 DUKE LAW & TECHNOLOGY REVIEW 85 (2017); Simon Stern, *Introduction: Artificial Intelligence, Technology, and the Law*, 68 UNIVERSITY OF TORONTO LAW JOURNAL 1 (2018); Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87 (2014); David C. Vladeck, *Machines without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117 (2014).

[2] For overviews and discussions of such applications, *see, e.g.*, Kevin D. Ashley, Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age (2017); Benjamin Alarie et al., *How Artificial Intelligence Will Affect the Practice of Law*, 68 University of Toronto Law Journal 106 (2018); Daniel Ben-Ari et al., *Artificial Intelligence in the Practice of Law: An Analysis and Proof of Concept Experiment*, 23 Rich. J.L. & Tech. 2 (2017); Kathryn Betts & Kyle Jaep, *The Dawn of Fully Automated Contract Drafting: Machine Learning Breathes New Life Into a Decades-Old Promise*, 15 Duke Law & Technology Review 216 (2017); Richard Berk & Jordan Hyatt, *Machine Learning Forecasts of Risk to Inform Sentencing Decisions*, 27 Fed. Sent'g Rep. 222 (2015); Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in e-Discovery Can Be More Effective and More Efficient than Exhaustive Manual Review*, 17 Richmond Journal of Law & Technology 1 (2011); David Lat, *How Artificial Intelligence Is Transforming Legal Research*, Above the Law (2018), https://abovethelaw.com/law2020/how-artificial-intelligence-is-transforming-legal-research/.

law.[3] Our interest in studying the adoption of machine learning tools in the legal profession was generated in part by a belief that lawyers—due to education and training, professional rules and ethical obligations, and their own interest in protecting themselves from professional liability—would have particular demands and expectations about the transparency, interpretability, explainability, and accountability of machine learning systems.

The concern that engineers and logics of automation will stealthily usurp or undermine the decision-making logics, values, and domain expertise of end-users has been an ongoing and legitimate complaint about decision support and other computer systems.[4] As technology reconfigures work practices, researchers have documented potential loss of human agency and skill[5] both over- and under-reliance on decision support systems,[6] confusion about responsibility,[7] and diminished accountability.[8] Scholars have also raised concerns with the entangled question of how technologically reconfigured

---

[3] Our research is ongoing. We are continuing to interview lawyers, in-house technical professionals, and legal technology company representatives.

[4] See Citron, Danielle Keats. "Technological due process." *Wash. UL Rev.* 85 (2007): 1249 (identifying the slippage and displacement of case worker values by engineering rules embedded in an expert system); Moor, James H. "What is computer ethics?" *Metaphilosophy* 16.4 (1985): 266-275 (identifying three ways invisible values manifest in technical systems—to hide immoral behavior, gap-filling during engineering that invisibly embeds coders' value choices, and through complex calculations that defy values analysis); Burrell, Jenna. "How the machine 'thinks': Understanding opacity in machine learning algorithms." *Big Data &* Society 3.1 (2016) (describing three forms of opacity in corporate or state secrecy, technical illiteracy, and complexity and scale of machine-learning algorithms); Pasquale, Frank A., "Professional Judgment in an Era of Artificial Intelligence and Machine Learning" (November 8, 2017), *Boundary 2* (forthcoming) (contrasting the reductionist epistemology and functionalist assumptions underlying substitutive automation with the holistic epistemology of professional judgment and the conflictual, political, and contestable nature of professional work, particularly in education and healthcare professionals).

[5] Lee J.D., Seppelt B.D. (2009) Human Factors in Automation Design. In Nof S. (eds) Springer Handbook of Automation. Berlin: Springer (detailing how automation that fails to attend to how it redefines and restructures tasks, and the behavioral, cognitive, and emotional responses of operators to these changes, produce various kinds of failure, including those that arise from deskilling due to reliance on automation).

[6] See Goddard, Kate, Abdul Roudsari, and Jeremy C. Wyatt. "Automation bias: a systematic review of frequency, effect mediators, and mitigators." *Journal of the American Medical Informatics Association* 19.1 (2011): 121-127 (reviewing literature on automation bias in health care clinical decision support systems)

[7] For an overview of research on technology-assisted decision making and responsibility, see Mosier, Kathleen L., and Ute M. Fischer. "Judgment and decision making by individuals and teams: issues, models, and applications." Reviews of human factors and ergonomics 6.1 (2010): 198-256 at 232-233.

[8] Nissenbaum, Helen. "Computing and accountability." Communications of the ACM 37.1 (1994): 72-81; Simon, Judith, "Distributed epistemic responsibility in a hyperconnected era." The Onlife Manifesto. Springer, Cham, 2015. 145-159

work practices shift power in ways that misalign with liability frameworks, leaving humans unable to exercise control but bearing the weight and blame for system failures.[9] For example, Elish (2019) explores how humans tend to take the brunt of failures in sociotechnical systems, acting as "moral crumple zones" by absorbing a disproportionate amount of responsibility and liability relative to their actual control and agency.[10]

In the context of automated decision making systems, scholars view increasing system interpretability and explainability as the primary strategy to ensure systems are aligned with domain-specific practices around the construction of knowledge and decision making and to address errors that stem from automation bias and self-reliance.[11] Research has, for example, examined the impact of various forms of explanatory material, including confidence scores and lists of important inputs (whether comprehensive or selective), on the accuracy of decisions, deviation from system recommendations, and trust.[12] But such research is mixed on the relationship between explanations and correct decision making.[13] The literature also generally assumes that correct prediction is the benchmark against which a machine learning system should be vetted. Yet, preferences for false positives and false negatives should inform evaluations of performance because different kinds of failures may be very differently weighted in different fields. For example, the legal system may have a preference for letting

[9] See, e.g., Jones, Meg Leta. "The ironies of automation law: tying policy knots with fair automation practices principles." Vand. J. Ent. & Tech. L. 18 (2015): 77.
[10] Elish, M. C., "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction (pre-print)" (March 1, 2019). *Engaging Science, Technology, and Society* (pre-print). Available at SSRN: http://dx.doi.org/10.2139/ssrn.2757236.
[11] *See* Nunes, Ingrid, and Dietmar Jannach. "A systematic review and taxonomy of explanations in decision support and recommender systems." User Modeling and User-Adapted Interaction 27.3-5 (2017): 393-444 (reviewing approaches to explanations in "advice-giving systems"; *see also* Bussone, Adrian, Simone Stumpf, and Dympna O'Sullivan. "The role of explanations on trust and reliance in clinical decision support systems." 2015 International Conference on Healthcare Informatics (ICHI), IEEE, 2015, pp. 160-169 at 160 (discussing research findings on automation bias and self-reliance)
[12] Bussone et al. (2015), supra note 9.
[13] *Id.* at 161 (describing different studies finding explanations leading to better and worse decisions).

guilty individuals escape punishment to protect against innocent individuals being incarcerated,[14] while medicine may have a varying preference for over- and under-identification of disease depending upon a range of contingent variables.[15]

That said, there are two notable gaps across the literatures investigating the use of machine-learning-based systems to aid in decision making. First, what are the roles of professionals, their organizations, and their professional environments in the adoption, implementation, and governance of such systems? How are these systems entering professional fields? What kinds of socio-material forces—for example, professional ethical duties, identity as an expert, the configuration of the system itself—affect how professionals understand and use such systems? Is the rise of automated decision making reconfiguring professional practices and the profession? If so, how? These questions have received limited attention in the literature.[16] Second, and relatedly, there is a *dearth of empirical data* on professionals, their organizational environments, and their interactions with today's automated, machine-learning-based decision making systems. Plenty of research in computer science and the FAT (Fairness, Accountability, and Transparency) subfield interrogates and evaluates the technical workings of such systems to shed light on such values as transparency, fairness, and accountability,[17] and an

---

[14] This preference in Anglo-American law is reflected in the so-called "Blackstone ratio:" [I]t is better that ten guilty persons escape than one innocent suffer." 4 William Blackstone 352, *Commentaries on the Laws of England* (1769). *See also* Richard H. Fallon, Jr., The Core of an Uneasy Case for Judicial Review, 121 Harv. L. Rev. 1693, 1706 (2008) ("[E]rrors that result in the conviction of the innocent are more morally disturbing than errors that result in acquittals of the guilty. In light of that assessment, we have adopted a system that minimizes the most morally grievous errors, even if that system leads to more of the less grievous errors, and indeed to more total errors, than would an alternative.")

[15] See Kraemer, Felicitas, Kees Van Overveld, and Martin Peterson. "Is there an ethics of algorithms?" *Ethics and Information Technology* 13.3 (2011): 251-260 (arguing that some algorithms are value-laden, illustrating with an exploration of how threshold choices for medical image may be dependent on context of use).

[16] Admittedly, this may, in part, be an artifact of researchers tending to focus on domains such as "predictive policing," risk-recidivism, facial recognition, ad targeting, or recommender systems, where professionals may play a limited role in the adoption, use, and governance of the systems.

[17] See, e.g., Alexandra Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, 5 BIG DATA 153 (2017); Amit Datta et al., *Discrimination in Online Advertising: A Multidisciplinary Inquiry*, CONFERENCE ON FAIRNESS, ACCOUNTABILITY AND TRANSPARENCY 20 (Jan. 21, 2018); Jon Kleinberg et al., *Inherent Trade-Offs in the Fair Determination of Risk Scores*, PROCEEDINGS OF THE

increasing amount of legal scholarship theorizes and makes normative claims about what laws or regulatory frameworks we need to address automated decision making systems.[18] But we also need rigorous, empirical social-science research into the professionals, their organizational environments, and the broader professional ecosystems in which these technical systems are embedded.

While traditional, rule-based expert systems have had a long history in professional fields like medicine, there are unique challenges posed by the use of today's predictive algorithmic systems, particularly machine-learning-based systems, to aid in professional decision making:

1) Whereas engineers of expert systems explicitly program in a set of rules, ideally based on the domain knowledge of adept subject matter experts, today's ML-based predictive algorithmic systems are designed—in effect—by deriving a set of decision rules from the data on which they train, which creates some unique challenges to ensuring systems accord with professional expertise and judgment.
2) Some of the algorithms used make it difficult to understand the rules they have learned from the data. Unlike an expert system where domain professionals can review and interrogate rules, these systems can provide insight into the inputs and outputs but lack the ability to easily interrogate the *rules or the reasoning by which the outputs were generated.*
3) The systems are dynamic, usually probabilistic, and they don't have one "right" answer. They need not make the same decision about two people (documents, etc.) the same way, nor must they make the same decision about the same person (document, etc.) at two points in time the same way. This plasticity challenges even the oversight provided by examinations of inputs and outputs.

8TH CONFERENCE ON INNOVATIONS IN THEORETICAL COMPUTER SCIENCE 43:1 (Berkeley, CA Jan. 9–11, 2017); Joshua A. Kroll et al., *Accountable Algorithms*, 165 UNIVERSITY OF PENNSYLVANIA LAW REVIEW 633 (2017).

[18] See, e.g., the many legal scholars and papers presented at the annual We Robot conference, among other gatherings (https://robots.law.miami.edu); ROBOT LAW (Ryan Calo et al. eds., 2016); Danielle Keats Citron & Frank A. Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASHINGTON LAW REVIEW 1 (2014); Andrew Tutt, *An FDA for Algorithms*, 69 ADMINISTRATIVE LAW REVIEW 83 (2017).

Thus, predictive algorithmic systems embed many subjective judgments on the part of system designers—for example, judgments about the training data, how to clean the data, how to weight different features, which algorithms to use, what information to emphasize or deemphasize, etc.

As discussed above, much of the recent scholarship on the automated decision support systems of today lacks rigorous empirical evidence regarding how those systems play out in professional practice. However, Christin's (2017) work on risk-recidivism and newsroom algorithms is a notable exception.[19] She takes an ethnographic approach to studying algorithms in action in expert fields[20]—namely, journalists and legal professionals working in the criminal justice system. Her work reveals that professional workers and managers appropriate machine-learning systems into their work practices, as they do other technology, informed by routines, norms, obligations of professional identity, and their position relative to others within an organizational hierarchy (e.g., managers vs. workers).

Similar to Christin (2017), we draw on the work of Science and Technology Studies (STS) scholars and organizational sociologists who study the socio-material interaction of technologies and

---

[19] Angèle Christin, *Algorithms in Practice: Comparing Web Journalism and Criminal Justice*, 4 BIG DATA & SOCIETY 1 (2017).

[20] Id. at 2. Christin (2017) distinguishes "expert fields" from professions (although they may overlap), defining expert fields as "configurations of actors and institutions sharing a belief in the legitimacy of specific forms of knowledge as a basis for intervention in public affairs." She makes this distinction for practical and strategic reasons. From a strategic standpoint, she makes the distinction in order to take a broader "field-based" analytical framework to her sociological analysis (drawing specifically on Pierre Bourdieu's conception of a "field"). We take an analogous, if conceptually distinct, systems-based view of the legal profession and lawyers, in which the profession is marked by constantly evolving processes of conflict, cooperation, and exchange with internal and external stakeholders. *See* ANDREW ABBOTT, THE SYSTEM OF PROFESSIONS: AN ESSAY ON THE DIVISION OF EXPERT LABOR (1988). And from a practical standpoint, she makes the distinction between expert fields and profession so that she can include in her comparative analysis journalists, who are not typically thought of as a highly professionalized occupation in the sense that, compared to highly professionalized domains like law, they lack, among other things, state-licensed monopoly control over the barriers to entry for their work. By staying within the legal profession to understand how lawyers are appropriating machine-learning decision support systems, we make no comparisons across professions here and thus see no need to follow Christin's distinction between "expert fields" and professions here. On Bourdieu's vision of "fields," *see* PIERRE BOURDIEU & LOÏC J.D. WACQUANT, AN INVITATION TO REFLEXIVE SOCIOLOGY (1992). On sociological field theory more generally, see Daniel N. Kluttz & Neil Fligstein, *Varieties of Sociological Field Theory*, *in* HANDBOOK OF CONTEMPORARY SOCIOLOGICAL THEORY 185 (Seth Abrutyn ed., Handbooks of Sociology and Social Research, 2016).

workers within organizational systems.[21] We also draw on insights from the sociology of the professions[22] and legal scholarship on the legal profession and technology.[23] Technology is only meaningful once humans engage with it in practice, but of course the material features and design of technology also exert strong influences on human perceptions and actions, and they can reconfigure organizational forms and work relations. In this way, the social and the technical are mutually constitutive.[24]

The mutually constitutive influences that the professionals can have on the adoption, use, and governance of predictive, machine-learning-based decision support systems (and vice-versa) is evident in the medical field, where clinical decision support systems are under scrutiny due to both explicit

---

[21] See, e.g., Stephen R. Barley, *Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organizational Studies*, 41 ADMINISTRATIVE SCIENCE QUARTERLY 404 (1996); Jeanette Blomberg et al., *Reflections on a Work-Oriented Design Project*, 11 HUMAN–COMPUTER INTERACTION 237 (1996); Christin (2017) (supra note 17); Wanda J. Orlikowski, *Sociomaterial Practices: Exploring Technology at Work*, 28 ORGANIZATION STUDIES 1435 (2007); Trevor J. Pinch & Wiebe E. Bijker, *The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other*, 14 SOC STUD SCI 399 (1984); Langdon Winner, *Do Artifacts Have Politics?*, 109 DAEDALUS 121 (1980).

[22] *See generally* ANDREW ABBOTT, THE SYSTEM OF PROFESSIONS: AN ESSAY ON THE DIVISION OF EXPERT LABOR (1988); MAGALI SARFATTI LARSON, THE RISE OF PROFESSIONALISM: A SOCIOLOGICAL ANALYSIS (1977); Sida Liu, *Boundaries and Professions: Toward a Processual Theory of Action*, 5 JOURNAL OF PROFESSIONS AND ORGANIZATION 45 (2018); Roy Suddaby & Daniel Muzio, *Theoretical Perspectives on the Professions*, *in* THE OXFORD HANDBOOK OF PROFESSIONAL SERVICE FIRMS 25 (Laura Empson et al. eds., 2015).

[23] *See, e.g.,* John Flood & Lachlan Robb, *Professions and Expertise: How Machine Learning and Blockchain Are Redesigning the Landscape of Professional Knowledge and Organization*, 73 UNIVERSITY OF MIAMI LAW REVIEW 443 (2019); A. Michael Froomkin et al., *When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced over-Reliance on Machine Learning*, 61 ARIZONA LAW REVIEW 33 (2019); Frank A. Pasquale, *Professional Judgment in an Era of Artificial Intelligence and Machine Learning*, 46 BOUNDARY 2 73 (2019); Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 IOWA L. REV. 1691 (2014); Dana A. Remus & Frank Levy, *Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law*, 30 GEO. J. LEGAL ETHICS 501 (2017); RICHARD SUSSKIND & DANIEL SUSSKIND, THE FUTURE OF THE PROFESSIONS: HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS (2015).

[24] On the mutually constitutive relationship between the technical and the social, especially with respect to work and organizations, see, e.g., Stephen R. Barley, *Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organizational Studies*, 41 ADMINISTRATIVE SCIENCE QUARTERLY 404 (1996); Wanda J. Orlikowski & Susan V. Scott, *Sociomateriality: Challenging the Separation of Technology, Work and Organization*, 2 ANNALS 433 (2008); Wanda J. Orlikowski, *Sociomaterial Practices: Exploring Technology at Work*, 28 ORGANIZATION STUDIES 1435 (2007); PAUL M. LEONARDI ET AL., MATERIALITY AND ORGANIZING: SOCIAL INTERACTION IN A TECHNOLOGICAL WORLD (2012).

regulatory oversight of medical devices *and* legal-ethical duties and perceptions of doctors themselves.[25] Both regulators[26] and doctors[27] are demanding that clinical decision support systems be aligned with the fields' decision-making processes, interpretable by medical professionals, and used under conditions that support "epistemically responsible" knowledge production and behavior.[28] Regardless of how exactly clinical decision support systems are regulated by the FDA, professional licensing requirements, ethical duties, and tort-based malpractice liability principles, and doctors' own conceptions of themselves as users of these technologies will shape clinical decision support tools *and* the conditions of their adoption and use. The exact contours of these various ethical and legal obligations are still emerging, but professionals and professional associations are keenly aware of the need to actively shape

---

[25] For our purposes, automated clinical decision support systems relying on machine learning to aid medical doctors in making decisions are a useful comparison for the predictive coding systems in law that we study below.

[26] In December 2016, President Obama signed into law the 21st Century Cures Act (Cures Act). Pub. L. No. 114-255 (2016). Section 3060(a) of the Cures Act added a new subsection to the Food, Drug and Cosmetic Act (FDCA) that excludes from the Food and Drug Administration's (FDA) medical-device regulations and approval processes "software function" that meets the following conditions: 1)not intended to acquire, process, or analyze a medical image or a signal from an in vitro diagnostic device or a pattern or signal from a signal acquisition system; 2) intended for the purpose of displaying, analyzing, or printing medical information about a patient or other medical information (such as peer-reviewed clinical studies and clinical practice guidelines); 3) intended for the purpose of supporting or providing recommendations to a health care professional about prevention, diagnosis, or treatment of a disease or condition; and, 4) intended for the purpose of enabling such health care professional to independently review the basis for such recommendations that such software presents so that it is not the intent that such health care professional rely primarily on any of such recommendations to make a clinical diagnosis or treatment decision regarding an individual patient. 37 21 U.S.C. § 360j(o)(1)(E)(i)-(iii) (2016)

[27] *See, e.g.,* statement of the American Medical Association, "Augmented intelligence in health care H-480.940." (Last modified 2018) (Accessed August 2018) (https://policysearch.ama-assn.org/policyfinder/detail/augmented%20intelligence?uri=%2FAMADoc%2FHOD.xml-H-480.940.xml) (setting out principles to guide development and use of AI); Evans, Emily L., and Danielle Whicher. "What Should Oversight of Clinical Decision Support Systems Look Like?." AMA journal of ethics 20.9 (2018): 857-863 (arguing that while using a clinical decision support system may not be a research activity under the Common Rule, its use requires more ethical and regulatory oversight than clinical practice and proposing a framework that sets out conditions governing use, ongoing monitoring of data quality, processes for developing and validating algorithms, and protections for patient data.); and for an example of how the professions is grappling with such machine learning decision support systems see, Martinez-Martin, Nicole, Laura B. Dunn, and Laura Weiss Roberts. "Is It Ethical to Use Prognostic Estimates from Machine Learning to Treat Psychosis?." *AMA journal of ethics* 20.9 (2018): 804-811.

[28] Judith Simon, *Distributed Epistemic Responsibility in a Hyperconnected Era*, *in* THE ONLIFE MANIFESTO: BEING HUMAN IN A HYPERCONNECTED ERA 145 (Luciano Floridi ed., 2015).

these tools to serve the needs of the medical field.[29]

There is a noted lack of empirical evidence on how lawyers think about discovery and go about conducting it in today's world of ESI and AI-based e-discovery systems.[30] The only empirical studies of decision support technologies and discovery practices at law firms that we have found were ethnographic studies undertaken in the 1990s by anthropologists and computer scientists tasked with designing systems to aid the litigation support team at a large law firm.[31] There, researchers found lawyers articulating and enacting a superior status relative to that of litigation support staff, whom they tended to see as only performing mundane, routine work of "document review" and incapable of the more complex decision making performed by attorneys.[32] Cautious and risk-averse, high status, and protective of professional expertise, lawyers were therefore reluctant to hand off anything beyond what they saw as routinized work. It is an empirical question as to whether and how that dynamic may be different with respect to attorneys and the workers who deal with machine-learning-based legal technologies today.

Our empirical research begins to fill this gap. We explore the adoption of machine learning decision support systems in the field of law. Our analysis of the relations among lawyers, litigation

---

[29] Char, Danton S., Nigam H. Shah, and David Magnus. "Implementing Machine Learning in Health Care — Addressing Ethical Challenges." *New England Journal of Medicine* 378, no. 11 (March 15, 2018): 981–83. https://doi.org/10/gddr8s.

[30] *See* Seth Katsuya Endo, *Discovery Hydraulics*, 52 UC DAVIS LAW REVIEW 1317, 1337 (2019) ("[T]here is little empirical data about what drives lawyers' choices in their discovery practices[.]"), citing Judith A. McKenna & Elizabeth C. Wiggins, Empirical Research on Civil Discovery, 39 B.C. L. REV. 785, 803 (1998) ("Much of the literature on incentives affecting discovery practice is rooted in economic theory. Yet, there is little information about how lawyers actually make discovery decisions.").

[31] Jeanette Blomberg et al., *Reflections on a Work-Oriented Design Project*, 11 HUMAN–COMPUTER INTERACTION 237 (1996) (describing their experiences designing a case-based prototype system for information retrieval and their observations of organizational politics and divisions of labor between attorneys and litigation support staff at the law firm while designing image-analysis technologies to aid in document review and classification); Lucy Suchman, *Working Relations of Technology Production and Use*, 2 COMPUT SUPPORTED COOP WORK 21 (1993) (arguing for industrial designers to be aware of the work practices of not only technology production but also its use among various users, and describing observations of the status hierarchies, contestable knowledge claims, and actions of lawyers and litigation support staff at a law firm).

[32] See Suchman (1994), supra note 35 at 32 (describing how litigation support work was invisible to attorneys and how attorneys described such work as a "mindless, routine form of labor").

support professionals (whether inside or outside the firm), and predictive coding technologies in the modern-day legal services field, provides a rich account of the actual effect of machine learning systems on legal practice, and identifies domain-specific challenges posed by current machine-learning-based systems and practices. The insights we offer open up new questions for the profession, and identifies new sites for interventions to shape the adoption, use, and governance of these tools going forward.

AUTOMATED LAWYERLY DECISION MAKING: TAR AND PREDICTIVE CODING FOR E-DISCOVERY

Our empirical case in this paper is lawyers' use of technology-assisted review (TAR), also called "predictive coding," systems for the discovery phase of litigation. One of the main challenges facing litigants today is the time and expense required to wade through ever-increasing amounts of electronically stored information (ESI) (e.g., data produced by smartphones, email, wearable devices, Internet of Things) during discovery. Lawyers and their clients must review their clients' records in order to search, collect, and produce those that are relevant and responsive to the other party's requests and not protected by legal privileges.

Discovery was an onerous process even in the days prior to ESI. With the vast amounts of ESI today, however, e-discovery can take up huge amounts of time for lawyers and clients tasked with manually reviewing ESI. It also entails huge monetary costs for litigants. A 2012 study by RAND researchers found that e-discovery production costs averaged about $18,000 per gigabyte of information, with costs attributable to document review being 70% or more of total e-discovery costs in more than half of the 57 cases studied.[33] The stakes of e-discovery are high in a different sense, with attorneys and clients exposed not only to the risk of adverse case outcomes but also to potential loss of attorney-client

[33] *See* Pace, Nicholas M., and Laura Zakaras. 2012. *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery*. Santa Monica, CA: RAND Corporation (analyzing the collection, processing, and review costs for e-discovery across 57 cases).

or other confidentiality privileges and even disciplinary action if they inadvertently produce or withhold otherwise discoverable ESI.

The concomitant rise of ESI and advancements in technology over the past two decades or so have spawned an ever-growing industry of e-discovery specialists, support staff, consultants, technology vendors, and products. Predictive coding systems, under the umbrella of TAR, are marketed as tools to aid legal professionals in managing, classifying, and reviewing ESI.[34] We focus on these e-discovery systems because, based on our conversations with lawyers and review of the literature, they represent one of the most well-developed applications of automated decision-support technology in the legal profession to date and because they provide a useful lens through which to discuss particular professional and ethical issues in their design and use.[35]

Broadly, TAR encompasses various technologies and techniques used on ESI, such as machine learning, clustering, semantic analysis, and sentiment analysis, to accomplish a broad range of tasks (e.g., email threading, de-duplication, document classification, visualization) that may or may not use predictive algorithms or machine-learning techniques to predict potentially responsive documents. Although most in the industry use "TAR" and "predictive coding" interchangeably, for simplicity, and because we want to focus attention on the machine-learning-based process of analyzing and predicting which documents among a corpus are responsive and not responsive during discovery, we will use "predictive coding" unless TAR is specifically used in quotes from the literature or our interviews.[36] The

---

[34] Examples of TAR products and ediscovery platforms on the market today include those from Brainspace (https://www.brainspace.com/), Catalyst (https://catalystsecure.com/), Exterro (https://www.exterro.com/e-discovery-software/data-management/predictive-intelligence/), H5 (https://www.h5.com/), Nuix-Ringtail (https://get.nuix.com/nuix-ringtail/), and Relativity (https://www.relativity.com/).
[35] Katie Shilton, *Values and Ethics in Human-Computer Interaction*, 12 HCI 107 (2018).
[36] Despite sharing some underlying general principles, predictive coding as used in the specific context of legal discovery should not be confused with predictive coding concepts and models as developed in neuroscience, cognitive science, and machine learning. For a review of predictive coding in these fields, *see* Huang, Yanping, and Rajesh P. N. Rao. 2011. "Predictive Coding." *Wiley Interdisciplinary Reviews: Cognitive Science* 2 (5): 580–93; *see also* Hinton, Geoffrey E. 2007. "Learning Multiple Layers of Representation." *Trends in Cognitive Sciences* 11 (10): 428–34; *see also* Clark, Andy. 2013. "Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioral and Brain Sciences* 36 (3): 181–204. For applications to signal

Electronic Discovery Reference Model (EDRM), a community of e-discovery and legal professionals housed at Duke University Law School that develops resources and best practices for information governance and e-discovery, provides two definitions of predictive coding:

> Predictive Coding
> 1. An industry-specific term generally used to describe a Technology-Assisted Review process involving the use of a Machine Learning Algorithm to distinguish Relevant from Non-Relevant Documents, based on a Subject Matter Expert's Coding of a Training Set of Documents. See Supervised Learning and Active Learning.
> 2. A group of machine learning technologies that predict which documents are and are not responsive based on the decisions applied by a subject matter expert to a small sample of documents.[37]

There are different machine-learning techniques used in predictive coding that require varying levels of human reviewer effort and varying degrees of initial training.[38] Cormack and Grossman have published a series of papers evaluating various predictive-coding systems against human reviewers and against each other.[39] They currently argue for continuous active learning (CAL), in which an attorney can continue to adjust the training algorithm during document review, as the best available method, although even they note the difficulty of defining, measuring, and achieving optimal reliability and

---

processing and data compression, *see* Shi, Yun Q., and Huifang Sun. 1999. *Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards*. Boca Raton, FL: CRC Press.

[37] EDRM, *Predictive Coding*, EDRM Glossary, https://www.edrm.net/glossary/predictive-coding/.

[38] *See* Cormack, Gordon V., and Maura R. Grossman. 2014. "Evaluation of Machine-Learning Protocols for Technology-Assisted Review in Electronic Discovery." In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, 153–162; *see also* Grossman, Maura R., Gordon V. Cormack, and Adam Roegiest. 2017. "Automatic and Semi-Automatic Document Selection for Technology-Assisted Review." In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, 905–908.*

[39] *See* Cormack, Gordon V., and Maura R. Grossman. 2014. "Evaluation of Machine-Learning Protocols for Technology-Assisted Review in Electronic Discovery." In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, 153–162; *see also* Cormack, Gordon V., and Maura R. Grossman. 2016. "Engineering Quality and Reliability in Technology-Assisted Review." In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 75–84. SIGIR '16. New York, NY: ACM; ———. 2016. "Scalability of Continuous Active Learning for Reliable High-Recall Text Classification." In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, 1039–1048.; ———. 2017. "Navigating Imprecision in Relevance Assessments on the Road to Total Recall: Roger and Me." In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 5–14.; Grossman, Maura R., Gordon V. Cormack, and Adam Roegiest. 2017. "Automatic and Semi-Automatic Document Selection for Technology-Assisted Review." In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 905–908.

recall while also limiting human reviewer effort.

Although predictive coding is usually not required in a case and instead often agreed to by the parties' attorneys, courts and the legal profession more generally have taken notice of its rise and have, on the whole, welcomed its use.[40] For example, judges have recently begun to specifically approve the use of predictive coding in published opinions.[41] The Federal Rules of Civil Procedure were amended in 2015 to include new rules to address e-discovery.[42] And professional bodies have convened to address

[40] That is not to say that all legal commentators have welcomed predictive coding with open arms. See, e.g., Seth Katsuya Endo, *Technological Opacity & Procedural Injustice*, 59 B.C. L. REV. 821 (2018) (arguing that the opacity of machine-learning-based predictive coding systems can undermine the due process of norm of participation, especially for parties who lack adequate understanding of the system's reasoning process); Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 IOWA L. REV. 1691 (2014) (recognizing predictive coding's potential benefits but cautioning that it also brings significant costs: 1) the tendency to overlook the wide variation in predictive coding systems' technical features and efficacy, 2) the erosion of lawyers' professional jurisdiction over discovery by delegating the process to non-lawyer computing systems, vendors, and technical specialists, and by lowering professional oversight standards, and 3) the undermining of client representation with threats to work-product and attorney-client privileges and confidentiality via new rules and norms pushing lawyers to cooperate with the opposing party by disclosing things like seed sets or system evaluation metrics).

[41] In 2012, US Magistrate Judge Andrew Peck became the first federal judge to publish an opinion approving the use of predicting coding software as an acceptable of means of conducting discovery. *See* Monique Da Silva Moore, et. al. v. Publicis Groupe & MSL Group, 287 F.R.D. 182, 193 (S.D.N.Y. Apr. 26, 2012) ("This Opinion appears to be the first in which a Court has approved of the use of computer-assisted review. That does not mean computer-assisted review must be used in all cases, or that the exact ESI protocol approved here will be appropriate in all future cases that utilize computer-assisted review . . . What the Bar should take away from this Opinion is that computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review."); *see also* Nat'l Day Laborer Org. Network v. U.S. Immigration & Customs Enforcement Agency (NDLON), 877 F. Supp. 2d 87, 109 (S.D.N.Y. 2012) ("[P]arties can (and frequently should) rely on . . . machine learning tools to find responsive documents."); Dynamo Holdings Ltd. P'ship v. Comm'r, 143 T.C. 183, 191-192 (T.C. 2014) ("Although predictive coding is a relatively new technique, and a technique that has yet to be sanctioned (let alone mentioned) by this Court in a published Opinion, the understanding of e-discovery and electronic media has advanced significantly in the last few years.... In fact, we understand that the technology industry now considers predictive coding to be widely accepted for limiting e-discovery to relevant documents and effecting discovery of ESI without an undue burden."); Rio Tinto PLC v. Vale S.A., 306 F.R.D. 125, 126 (S.D.N.Y. 2015) (holding that technology-assisted review [TAR] is "an acceptable way to search for relevant ESI in appropriate cases.").

[42] *See, e.g.,* the recently revised Rule 26 of the Federal Rules of Civil Procedure, which governs discovery in civil litigations (amendments effective December 2015); *see also* Serhan, Stephanie. 2017. "Calling an End to Culling: Predictive Coding and the New Federal Rules of Civil Procedure." *Richmond Journal of Law & Technology* 23 (2): 1–36 (reviewing the 2015 amendments to the Federal Rules of Civil Procedure as applied to predictive coding, the split at that time among courts over when to use predictive coding during a case, and arguing that predictive coding should be done at the outset of discovery on the entire set of ESI rather than an already-culled set of documents); *see also* Schieneman, Karl, and Thomas C. III Gricks. 2014. "Implications of Rule 26(g) on the Use of Technology-Assisted Review." *Federal Courts Law Review* 7: 247–84 (evaluating the proper way to use

the evolving landscape of e-discovery technologies and issue best practices.[43]

Legal framework for responsible, ethical use of predictive coding

The governance of the legal profession's use of predictive coding is based in professional normative principles of attorney ethical duties of responsible conduct and competent representation of clients. Thus, to the extent that there are formal rules over predictive coding, they tend to be formalized in jurisdiction-specific ethical guidelines, model rules of professional conduct, and case law.[44] The American Bar Association's Model Rule 1.1 of the Model Rules of Professional Conduct states: "A lawyer shall provide competent representation to a client. Competent representation requires the legal knowledge, skill, thoroughness and preparation reasonably necessary for representation."[45] In 2012, the legal profession began the process of establishing a legal duty of *technological* competence on lawyers when the ABA's House of Delegates amended Comment 8 to Model Rule 1.1 to read:

> To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology, engage in continuing study and education and comply with all continuing legal education requirements to which the lawyer is subject.[46]

As of the end of February 2019, 36 states have formally adopted the amended comment to Rule

---

TAR at different phases of e-discovery for purposes of the attorney's reasonable-inquiry and certification requirements under Rule 26(g)).

[43] *See, e.g.*, The Sedona Conference, *The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 Sedona Conf. J. 1 (2018); EDRM, Frameworks and Standards: Technology Assisted Review (2018), http://www.edrm.net/frameworks-and-standards/technology-assisted-review/.

[44] We focus on formal governance mechanisms here—e.g., ethical duties instantiated in rules or laws, backed by the sanctioning power of courts or regulatory bodies. Although our respondents may mention informal guidelines or rules of thumb, such as best practices developed by professional groups (e.g., Sedona Conference Working Groups on e-discovery issues (Working Groups 1-2, 6-7) and the EDRM at Duke Law School) (supra note 41) or in-house discovery protocols developed by vendors/consultants or attorneys themselves, they invariably tell us that the foundations and guideposts for their informal governance mechanisms are rooted in the more formal rules of civil procedure regarding discovery and rules of ethical and professional conduct discussed here.

[45] The ABA Model Rules of Professional Conduct can be found at https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/model_rules_of_professional_conduct_table_of_contents.html.

[46] *See generally* § 15:5 of Grenig, Jay E., and William C. Gleisner. 2017. *eDiscovery and Digital Evidence*. Eagan, MN: Thomson West (reviewing and explaining relevant legal standards and ethical duties regarding attorney technological competence in e-discovery).

1.1.[47] On February 26, 2019, Texas became the most recent state to adopt the ABA's Comment 8 to Rule 1.1, when the Supreme Court of Texas amended Paragraph 8 of the comment to Rule 1.01 of the Texas Disciplinary Rules of Professional Conduct comment to track the ABA's model language.[48]

Although California has not specifically adopted the language of the ABA's Comment 8 to Rule 1.1 into its own rule of professional conduct regarding competency,[49] the State Bar of California has, since 2015, nevertheless incorporated the model rule's duty of technology competence with respect to e-discovery via a formal ethics opinion.[50] This opinion is particularly instructive not only because California is home to a thriving technology sector but also because it provides an extended discussion of attorney competence specifically as applied to conducting e-discovery during litigation. Attorneys should be able to perform the following nine skills:

> 1) initially assess e-discovery needs and issues, if any;
> 2) implement/cause to implement appropriate ESI preservation procedures;
> 3) analyze and understand a client's ESI systems and storage;
> 4) advise the client on available options for collection and preservation of ESI;
> 5) identify custodians of potentially relevant ESI;
> 6) engage in competent and meaningful meet and confer with opposing counsel

---

[47] Robert Ambrogi, *36 States Have Adopted Ethical Duty of Technology Competence*, Robert Ambrogi's LawSites (March 12, 2019), https://www.lawsitesblog.com/tech-competence/ (providing running tally of states that have adopted the ABA's comment to Model Rule 1.1 and links to each state's rule).

[48] *See* Supreme Court of Tex., Order Amending Comment to the Texas Disciplinary Rules of Professional Conduct, Misc. Docket No. 19-9016 (Feb. 26, 2019), available at http://www.txcourts.gov/media/1443638/199016.pdf (last visited March 20, 2019) ("Because of the vital role of lawyers in the legal process, each lawyer should strive to become and remain proficient and competent in the practice of law, *including the benefits and risks associated with relevant technology*") (amended language in italics)).

[49] California's professional rule regarding attorney competence is Rule 3-110 of the Rules of Professional Conduct of the State Bar of California. It holds:
(A) A member shall not intentionally, recklessly, or repeatedly fail to perform legal services with competence.
(B) For purposes of this rule, "competence" in any legal service shall mean to apply the 1) diligence, 2) learning and skill, and 3) mental, emotional, and physical ability reasonably necessary for the performance of such service.
(C) If a member does not have sufficient learning and skill when the legal service is undertaken, the member may nonetheless perform such services competently by 1) associating with or, where appropriate, professionally consulting another lawyer reasonably believed to be competent, or 2) by acquiring sufficient learning and skill before performance is required. (https://www.calbar.ca.gov/Attorneys/Conduct-Discipline/Rules/Rules-of-Professional-Conduct/Current-Rules/Rule-3-110) (last visited March 21, 2019).

[50] Cal. St. Bar Standing Comm. on Prof'l Resp. & Cond., Formal Op. No. 2015-193 (June 30, 2015) at 3 (quoting the revised Comment 8 to ABA Model Rule 1.1 to state that "[m]aintaining learning and skill consistent with an attorney's duty of competence includes keeping 'abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology[.]'")

concerning an e-discovery plan;
7) perform data searches;
8) collect responsive ESI in a manner that preserves the integrity of that ESI; and
9) produce responsive non-privileged ESI in a recognized and appropriate manner.[51]

In California, as in other states adopting the revised comment to the ABA model rule, if an attorney does not possess the requisite skills described above, they can satisfy their ethical obligation of e-discovery technology by associating with competent co-counsel or expert consultants. Such an expert could be a vendor, a subordinate attorney, or even the client itself, as long as they possess the necessary expertise.[52] However, associating with an expert raises another ethical duty—the duty to supervise—and potential tensions regarding professional expertise between attorneys, technology vendors, and clients that we address in our findings. Even if an attorney associates with a co-counsel or e-discovery consultant with expertise in handling e-discovery technology, that attorney, still has the responsibility to supervise such an expert and is ultimately responsible for the work of the expert:

> This consultation or association [with an expert], however, does not absolve an attorney's obligation to supervise the work of the expert under rule 3-110 [California's rule of professional conduct regarding competency], which is a non-delegable duty belonging to the attorney who is counsel in the litigation, and who remains the one primarily answerable to the court. An attorney must maintain overall responsibility for the work of the expert he or she chooses, even if that expert is the client or someone employed by the client. The attorney must do so by remaining regularly engaged in the expert's work, by educating everyone involved in the e-discovery workup about the legal issues in the case, the factual matters impacting discovery, including witnesses and key evidentiary issues, the obligations around discovery imposed by the law or by the court, and of any relevant risks associated with the e-discovery tasks at hand.[53]

RESEARCH DESIGN

We draw primarily on qualitative evidence obtained from approximately 26 hours of semi-structured, in-depth interviews of 25 respondents who work with predictive systems in the legal

---

[51] *Id.* at 3-4.
[52] *Id.* at 5.
[53] *Id.* at 5. The rule governing attorney competence in California, cited within this quote from the 2015 ethics opinion as Rule 3-110 of the Rules of Professional Conduct of the State Bar of California, is now cited as CA ST RPC Rule 1.1 (Business and Professions Code section 6068(e)) (new rules approved by the Supreme Court of CA May 10, 2018, effective Nov. 1, 2018).

profession—attorneys, litigation support staff working in law firms, and managers at companies that provide decision-support technology products and services to lawyers.[54] We recruited respondents via email and phone with a prepared recruitment script. We used snowball sampling techniques, with initial leads generated from our professional contacts in law and the legal technology industry and later leads gathered from interview respondents themselves. We audio-recorded interviews with participants' informed consent. Interviews lasted between 46 minutes and 2 hours, with an average length of 1 hour 2 minutes.

After transcribing the audio recordings, we used qualitative data analysis software MAXQDA to code the interview data according to categories of themes. To do this, we relied on a combination of inductive reasoning, which allowed for emergent issues and topics, and deductive reasoning, which was informed by our prior research and past experience as attorneys ourselves. In order to protect confidentiality, we use pseudonyms when reporting any names of interview respondents and their affiliated firms or companies.[55]

Of our 25 respondents, 17 work at law firms (12 attorneys, 5 litigation/technical support staff at law firms) and 8 work at legal technology companies. Within this latter group, all of their positions are at the management level: CEO (2), CTO (2), COO (1), Vice President (1), Director of Consulting (1), and Litigation Manager (1).[56] All respondents are based in the United States, although their firms/organizations do business overseas, as well.

Our sample of attorneys is not representative of the population of attorneys in the US, of course. For example, attorneys in our sample all work at law firms with greater than 50 attorneys and, with the exception of one respondent attorney working at a large plaintiff-oriented firm, would be roughly

[54] This research is ongoing, so numbers, and findings may change.
[55] Readers should infer no connection to actual names of respondents, their organizations, or any other actual person or organization.
[56] Three of these respondents are also founders of their companies (2 Founder/CEO; 1 Founder/CTO).

classified as corporate defense law firms. Because we focus on decision-support tools applied to the e-discovery context, all law firms represented have significant litigation practices. Our focus on attorneys and legal tech managers working with these kinds of law firms was strategic, as our research indicates that they are the firms most likely to be targeted as potential customers by technology company vendors and are the firms mostly likely to have the clients, resources, and types of cases (i.e., cases with significantly large volumes of electronically stored information (ESI)) that call for the use of automated decision-support technologies for e-discovery, such as predictive coding tools. In other words, they are the firms and attorneys most likely to have experience with and knowledge of these technologies. Thus, to the extent that our data identify challenges posed by the introduction and use of such systems, any conclusions we draw are likely to be conservative, if anything. As the project continues, we will supplement our sample of legal professionals with key informants working as attorneys at plaintiff-oriented law firms and as judges.

RESULTS

The rise of TAR and predictive coding in law firms and the legal profession

To what do lawyers and legal professionals in the surrounding legal services environment attribute the rise of predictive coding systems, and TAR more generally, in the legal profession?

Cost-cutting

Our respondents consistently positioned TAR as a cost-cutting strategy. Like the well-established practice of outsourcing to contract attorneys, and to out-of-country attorneys, delegating “document review” tasks to technical tools is first and foremost viewed as a tool for reducing litigation costs. The rising costs of litigation were a product of both escalating lawyer fees and the explosion in documents produced by daily corporate activities in the digital age. Our respondents viewed TAR and predictive coding primarily as a response to this unprecedented growth in ESI. As Carrie Lewis, a partner at a law firm, explained:

> We're seeing more and more that the general counsel has to show to their leadership and to their board that they have reduced costs by X percent or increased the use of technology. Then they're coming to us and saying how do we measure this? How do we show this? What do we do? (Carrie Lewis, attorney)[57]

Jason Ellison, currently a manager at a vendor and formerly a litigation support specialist within a law firm, echoed that sentiment: "Clients are increasingly looking at their spends. They're increasingly analyzing line items on bills and pushing down on law firm clients. This is something that started to get a lot of attention about ten years ago."

The increased reliance on technology to facilitate document review, combined with the growing practice of procuring legal services,[58] has fueled the growth of legal support platforms that manage and secure the voluminous corpus of documents generated by a business' general operation, as well as provide document review services.

Our interviewees explained that the evolution of TAR tools and pressure to cut costs is reorganizing the relationship between lawyers, vendors, and clients. This reorganization takes several forms. As to payment, most commonly, a corporate client (i.e., a party to the litigation) pays the e-discovery vendor, but the day-to-day interactions take place between the vendor and the attorneys. As Ken Summers, an executive at a TAR vendor, explained: "The vast majority of our clients are corporations typically in the Fortune 200…in 90% plus of the cases…[they]…pay us. They're the true client. The client we work with on a day-to-day basis is the law firm that represents those corporate clients. That's the structure."

Increasingly, it seems from our respondents, legal tech vendors are also marketing their services not only to law firms but directly to the corporations. One respondent, who wished not to be recorded

---

[57] Samir Anand, another attorney respondent, spoke to the issue of lawyer fees: "It is common knowledge now from judges that everyone's using software . . . . What's happened is lawyer rates have gone up so high that everyone just assumes that it's [predictive coding] being done by somebody else." (Samir Anand, attorney)

[58] Silvia Hodges Silverstein, *What We Know and Need to Know about Legal Procurement*, 67 S. C. L. Rev. 485 (2016).

but had managed an overseas office of e-discovery technicians and document reviewers for an e-discovery vendor, described her experience of being told by her supervisor to skip going through the law firm with which her company had contracted and go straight to the corporate client and attempt to cultivate a direct relationship with the corporation in the hopes of future work.

The increase of technically sophisticated vendors offering complex computational systems to aid legal decision making, combined with the pressure to reduce legal costs, has also led to reconfigurations in the extent to which law firms have control over the technologies that they use for their own work. Paul Young, an executive at an e-discovery vendor, explained:

> I can say with a fair degree of accuracy that in the industry in general…it is trending towards the corporate and that is a byproduct of, I believe, cost-cutting….there's a lot of corporations that…don't want to necessarily be beholden to law-firm direction and guidance. It's actually related to spending money. I don't really know if I can put it more gracefully than that, but the reality is if you're a multi-billion dollar company, do you really want a law firm that charges $1,000 an hour making all of your decisions for you? Or do you want to have people internally [who are] definitely looking out for your best interests and vetting outsourced vendors accordingly, contracting directly with them, managing that process internally versus going through a law firm? (Paul Young, e-discovery vendor executive)

This seems to be the case particularly with larger corporate clients. As one partner at a law firm told us:

> The overwhelming trend is that the lawyers are being taken out of that process…decisions about which corporate lawyers to use have been centralized by the client—which time-entry programs to use, which billing software to use, which ways that we report to the client—are all governed by terms and conditions from the client at the beginning of the relationship . . . . If you're looking especially at large, institutional clients…[t]hey want to make staffing decisions; they want to make the decisions on actually who should be reviewing documents and all of that . . . . This decision of who's reviewing documents and all that never even comes to me. (Samir Anand, attorney)

Angus Martin, a partner at a law firm, focused instead on the preferred-provider aspect of his dealings with large corporate clients and technical systems:

> The client will say—and it tends to be the Fortune 500 client—will say, 'We have a contract to do all of our e-discovery litigation with XYZ vendor.' It means that they get a better price on it. XYZ knows their data systems better, so they [the client] don't need to go out and pay my hourly rates [for me] to go learn how their servers are set up and all that kind of stuff. (Angus Martin, attorney)

It also appears that larger firms are using vendor platforms to further reduce costs and uncertainties of litigation through longer-term arrangements, standardization across litigation matters, and use of broader information-governance services that integrate litigation support. Echoing what other respondents told us, Egon Graham, a director at a vendor offering e-discovery consulting, explained that her company now provides a wider array of information-governance services beyond e-discovery, including "development of data policies, so everything from mobile devices, social media, definitely records-retention and disposition schedules. We work on implementing those. We consult on privacy . . . And we also do e-discovery playbooks—so making sure they are ready in the event they have discovery."

Even when vendors are hired by law firms, however, they are providing far more than a technical system with implications for lawyer-client relationships. Some of the larger vendors of these technical platforms are actually offering a mixed system of technical tools and humans, as Graham went on to explain:

> We have an array of products, and the client will tell us what review platform they want it to go up in. Then they will tell us if they want to do the whole review themselves. If they want just staffing, they just want some attorneys, we have a staffing arm…so we can give them just bodies to do review. If they want us to actually run their review for them, then we have a managed review set that will set up the workflows, do all the batching of documents, do the quality controls, give reports back, so they can set that up for them. Then we have production environments as well where we can help them produce the documents. (Egon Graham, director, e-discovery vendor)

Consulting staff at the vendors often includes a range of other experts, including statisticians, linguists, and data scientists, who play an important role in how predictive coding tools are used and interpreted in the discovery process. For example, one vendor representative explained their business and staffing to us:

> …our model is that we do not sell AI tools. We sell AI as a service. When corporate clients come to us, they will either provide the document analysis of key document identification, having it performed by attorneys who use AI tools, or they will purchase the service that [we] provide, which really is a combination of advanced technologies. The main difference is the technologies are applied by computational linguists and computer scientists who operate these technologies in a somewhat

different way than lawyers would. (Ken Summers, vendor executive)

Improved performance

While cost savings and the steep increase in the volume of material in discovery proceedings appear to be the key drivers of TAR, we did encounter the standard refrain of Big Data and machine learning advocates that algorithmic systems are better—less biased, more consistent and predictable—than fallible, sometimes malicious, humans. And this sentiment came not only from tech company representatives but from within the law firms. For example, Joe Goodman, who is a law-firm litigation support manager and works closely with attorneys at his firm, said:

> Yes, unequivocally, [TAR is] generally considered to be more accurate because it's an algorithm. It's not a human who blinked at the wrong time or got distracted by their dog or a search term was wrong and pulling back the wrong data, those kinds of things. There's so many reasons why a human review is flawed compared to using the technology." (Joe Goodman, litigation support manager at law firm)

Even when asked about training data and other variables that could influence model performance, Goodman did not waiver in his assessment of the relative accuracy of predictive coding tools compared to humans: "Those factors don't really play into it. It's a matter of comparing like populations, or two identical populations, for human review versus algorithmic review. You're going to see greater accuracy from the algorithmic review almost every time than you would from humans."

What did respondents have to say about human review of system performance?[59] Lawyers did report using human review as check on system outputs. However, its use was selective in ways that, if typical of practice, risks introducing a systematic bias of under-disclosure. For example, Goodman, the litigation support manager at a law firm discussed above, prefaced his discussion about lawyers' interaction with TAR systems with "[i]t's funny, it's almost always driven by volume." He went on to

[59] Compared to lawyers, vendors reported more reliance on, and evinced much deeper understanding of, traditional model-evaluation metrics like recall and precision. With respect to human review, they did not have objections to it, but on this, they tended to recommend the minimal amount of human review that, in their analysis, would best balance satisfying defensibility standards from the courts and save costs on human review for their clients.

explain how during the early stage of discovery, if his team is developing search terms to set the initial training set of documents, attorneys will say "… 'Here, run this group of search terms,' and…want to know how many documents it brings back. Then they say, 'Oh, that's too many. We've got to change the terms.' They've set the terms based on the number of documents that they've returned. Then they get the other side to agree to the search terms we're using and vice-versa." He understood that this was no proper way to determine responsiveness or address the discovery principles of proportionality and defensibility: "That's usually how that goes. It's very funny, and I've never really understood this. How is it that we're determining what to review based on how many documents come back on a given search term set? Either the search terms are perpetually responsive or they're not." When asked to explain the lawyers' reasoning he said, "They say 'Yeah, no, that's way too many documents. Let's use different terms.'" He assumed this approach by the lawyers is based on "(c)ost, effort, and time."

For their part, lawyers indicated being particularly averse to certain kinds of failures, namely the inadvertent production of privileged material. This leads post-predictive coding human reviews by attorneys to focus on documents that the system identified for production (i.e., documents scored by the predictive coding system at a probability that meets or exceeds the system's decision threshold to be classified as responsive, notwithstanding any other privileges or exceptions that might prevent disclosure). It also leads the human reviewers to give comparatively less attention, if any, to those documents *not* classified as potentially responsive (i.e., predicted negatives) or to conduct a systematic review of both groups of scored documents (i.e., both the predicted positives and predicted negatives). Our respondents reported very little questioning or real review of TAR model performance with respect to false negatives (i.e., documents that are actually responsive but not classified as such by the predictive system).

Implications: Ethics and values

Our interviews raise important issues with respect to professional ethics and values, the exercise of professional judgment, and the practice of law. Here, we focus specifically on the duty of competent representation and the attendant issues it raises when considered in light of predictive coding technologies and lawyers.

Duty of competent representation

As discussed earlier in the paper, attorneys have a professional ethical obligation to provide competent legal representation to their clients. For attorneys in most states today, that duty of competence entails keeping abreast of changes in the law and its practice, "including the benefits and risks associated with relevant technology."[60] What is happening in practice? And what do lawyers and legal technology professionals have to say about technology competence with respect to automated decision support systems, particularly when it comes to TAR and predictive coding systems?

First, and considering the issue of technical expertise before getting to the more specific issue of lawyers' ethical duty of technology competence, our TAR vendor respondents felt strongly that, compared to lawyers, they have the most technical expertise regarding information retrieval and predictive coding systems. As one executive-level manager explained to us:

> [T]his is a distinct professional domain, information retrieval. . . . It's truly a distinct professional field. I don't believe that at scale any company or any law firm or a company like ours can have truly two completely distinct twin core competencies. [Company name], I think, is probably today the best information retrieval company in the known universe when it comes to data analytics and litigation, investigations, etc., but we will never be a great law firm, even if we tried. It's just two distinct professional domains. Great law firms will never be great medical centers or medical practices. They'll never be great airlines, and they won't be ever great information retrieval companies. It just will not happen. These are two distinct professional companies. (Ken Summers, vendor executive)

Later in the interview, Summers described the average lawyer as "a lay person" who is ill-equipped to leverage "the scientific domain of search and review and information retrieval," leading to

---

[60] As we discussed above, 36 states have now formally adopted the American Bar Association's 2012 revised Comment 8 to Rule 1.1 of the Model Rules of Professional Conduct. *See* Ambrogi, supra note 47. Other states, such as California, can impose the same or similar duty through state bar ethics opinions. *See* State Bar of California Formal Opinion No. 2015-193, supra note 50.

"inefficiencies."

Matt Rogers, an attorney, expressed the general concerns that lawyers may have about responsibility: "Where are the responsibilities if the platform gets screwed up? Or you make mistakes? Or you make a representation that's belied by the data? That kind of thing." He went on to observe that the new arrangements between attorneys, clients, and technical expert vendors produced "decision-making friction…between what a [firm] wants to do, and what a client wants to do, and what the third-party provider wants to do."

Nevertheless, going against the stereotype of risk-averse lawyers and our expectation that our lawyers would point to concerns about liability risk due to inadequate understanding of the black-box nature of machine-learning-based tools, our interviews indicated an overall *lack* of concern about potential professional malpractice liability risk when discussing their views on the factors driving adoption and use of automated systems for lawyer decision making. For example, one attorney at a large firm who oversees the procurement of technical systems for the firm's lawyers and interacts frequently with attorneys and vendors providing these systems, explained when describing an automated contract review technology in use at the firm:

> The attorneys at [firm] are so diligent and so focused on providing value to their clients that—well, the best legal services for the client, even if it's not value in terms of dollars and cents—no one's been worried that this is going to be a shortcut that leads to some sort of malpractice problem. (Chad Mankins, attorney)

Responsibility and the duty to supervise

Even if a lawyer lacks an adequate understanding of the algorithms and models underlying TAR and predictive coding, as we discussed earlier, he or she can satisfy the ethical obligation of competent representation by associating with and supervising a sufficiently competent lawyer (within or outside the firm) and even a non-lawyer technical expert.[61] Given the deep technical knowledge of some of our in-

[61] See, e.g., the discussion of California's rules governing associating with co-counsel or technical experts for purposes of satisfying the duty of competent representation, supra note 49.

house litigation support staff respondents and consultants from predictive coding vendors whom we interviewed, we are confident that they would be deemed to be experts for these purposes. However, the primary attorney on the case still must exercise supervision over the work of the expert(s) employed and is ultimately responsible for legal representations made to opposing counsel and the court. How are lawyers and predictive coding vendors thinking about these issues of technology competence and the duty to supervise?

Egon Graham, who is a licensed attorney but works for a TAR vendor as an e-discovery consultant, had a particularly illuminating response when asked about the extent to which he "owns the discovery process" (his phrase) in his role as consultant:

> The rules of professional responsibility say that attorneys must supervise non-attorneys, and they also say that attorneys cannot practice for any company that is not wholly owned by attorneys. My company is not wholly owned by attorneys, which means we are not a law firm, and it therefore means I can't practice law. Even though I and others on my team have decades of experience among us and we're all licensed in multiple states, we cannot practice law as we sit here as serving consultants. We can consult. We have to be supervised by an attorney. It means I can't just have a paralegal hire me, and there's no one else that my stuff is going through for the discovery side of things.
>
> Whatever we do, we can, basically, own as much as the process as our client wants us to and control it as if we were the attorneys. We have all run these types of cases when we were practicing attorneys. At the end of the day, I need to be disclosing everything that I'm doing to an attorney so that they can satisfy their duty to supervise me, and as a barred attorney myself, it creates this weird duty to be supervised. Unlike any expert where it's not [pause]—the attorney has to understand everything I've done. They have to make sure that I'm not clearly just being reckless and doing things I shouldn't do, and if there's a big decision to be made, consulting with my client, making sure they're educated around their different options, and making a recommendation to them. (Egon Graham, vendor consultant)

Graham's response reveals some of the tensions and that the duty to supervise imposes on the attorney-client-vendor relationship in practice. As a licensed attorney himself, he exhibits a keen awareness of the ethical requirements and the limitations they impose on his interactions with the supervising attorney (e.g., be hired by the attorney, keep the attorney informed). Yet, he also characterizes himself as a deep expert (e.g., "decades of experience") and able to control the e-discovery process "as if we were attorneys."

Reflecting on this shift to technology vendors, and vendors' crucial (but often downplayed) human workers who manage discovery on a day-to-day basis, most lawyer respondents were clear that the lawyer remained responsible for mistakes in production, even if the mistakes were made by the vendor. Matt Simpson, a managing project attorney at a law firm, exemplified this sentiment: "[Y]ou cannot outsource that [responsibility] to the vendor. I know the vendors can really be helpful with the consulting work, but the lapels that the client and the court is going to grab is the firm." Similarly, Samir Anand, partner at a law firm, said: "… [M]y guess is all lawyers know that they are the ones responsible. I mean, none of us go to court and say, 'well, we used e-discovery software, so that was the problem.'"

The conviction with which our lawyers knew that the managing attorney is ultimately responsible for e-discovery does not take away from the fact that, in practice, things may get murkier. Clay Simpson, in-house managing project attorney at a law firm, voiced concern about law firms (not his own) where "responsibility is being outsourced to the vendor:"

> Q: What's your sense of the field in general, the legal field, about knowledge of those duties, commitments, or actual practices to have those clearly defined accountability chains?
>
> A: Yeah, I'd say it's hit or miss. I'd say it's limited. A lot of folks will – a lot of other law firms will have maybe an e-discovery practice, but they're not looped in on these issues. I think it varies widely. I made the outsource-to-the-vendor point because that's what I hear a lot of my peers complain about. A lot of that responsibility is being outsourced to the vendor. It's all great if everything works perfectly. Oftentimes, a good vendor consultant can testify, even if you're being challenged, but I think if something goes really poorly, I think that's probably a bad strategy.
>
> Q: What do the vendors have to say about the issue? Is that something that they're worrying about as a liability issue?
>
> A: They're worried about it; there's no doubt about it. You may talk to a vendor who's really invested in a great consultant to try to navigate those murky issues. I've talked to vendors just recently at conferences that have actually dealt with this issue. I had somebody come up to me after [a conference] and said, 'I ran into the same thing where it got outsourced to me. The person there was trying to combat it and push it back to the firm, then things went poorly, and then it was just a 'ball dropped' type of scenario. That's rough if the law firm itself isn't owning it, owning the project

> management and the AI side of it. (Clay Simpson, in-house managing project attorney)

Vendors even invoke specific strategies to protect themselves from blame and to keep firm lawyers responsible for legal determinations. For example, Paul Young, manager at an e-discovery vendor told us that his company "tr[ies] to avoid definitively saying anything like, 'you can stop now,' or 'this is definitely good enough.'" He explained that "that's legal determination," "we're not the outside counsel," and "[o]ur job is to arm outside counsel with all the info they need in order to make that determination."

Finally, although it was not a focus of interviews, there is a question about the extent to which attorneys may confuse contractual obligations with professional obligations. For example, when asked about whether lawyers have an ethical duty to inform clients about the use of a given technology product, one in-house litigation support manager suggested that the contractual arrangement between corporate client and vendor addressed this concern:

> …whenever we [law firm] contract with vendors, typically we get the client to sign the letter of engagement with the vendor directly so we don't act [as] the middle man for payment. We want the client to be on the hook to pay the vendor directly so we're out of the loop on that. They know what they're getting into. They know what they're signing up for. They know what tools are going to be used and they'll know how much it's going to cost, and they're in agreement with those terms. (Joe Goodman, litigation support director at law firm)

This assumption that service procurement will address concerns about whether or not technical choices should be discussed specifically with the client points to the risk of confusion about who is accountable for what in this newly triangulated relationship.

Interactions with opposing counsel

Finally, we address a different aspect of ethical duties that bears on attorney competence: how attorneys interact with opposing counsel during discovery. All attorney respondents expressed a preference for working with the other side to agree on the use of predictive coding. Matt Rogers, an experienced partner who heads the e-discovery practice at his firm, reflected this preference:

> If they're looking at – you throw them your non-responses, and they say, 'Hey, this is just not – you're not picking up a certain issue.' 'Sorry about that, we'll pick that up.' 'Your precision is, at this level, we would like it to be higher.' Maybe you agree beforehand on what it is. If the parties are being cooperative, it can be very productive, actually, to get people – I mean, you're holding down costs on both sides. (Matt Rogers, attorney)

However, our attorney respondents were not particularly worried about learning everything they could learn about their adversary's predictive coding system (e.g., seed-set disclosure, scoring/ranking methods, evaluation metrics), relying instead on their own expertise, following guidance from their own e-discovery vendors, and expressing trust in their opposing counsel not to act nefariously.[62] As Paul Young, a manager at an e-discovery vendor told us, he had "never been asked to explain why certain data subsets were not produced by virtue of some cutoff that left them out of the production universe." This could be due to his company being "proactive" and developing comprehensive defensibility plans for their clients, as he suggested, but it could also point to trust and lack of skepticism from opposing counsel.

Finally, speaking to issues of competence as well as reflecting the preference for agreements between the parties to reduce any potential ethical violations, Robert Baker, attorney at a large defense firm, stated:

> I don't think that we would ever run any kind of software and rely on it without really understanding what it was doing and the potential downsides would be. I think that's just for all the reasons from earlier. It's a due diligence question, basically. I think that having a stipulation from the other side is pretty close to a proxy for confidence. I think both sides agree to something, it's difficult for both sides to be incompetent at once, I think. It happens but just in the sense of if both parties are agreeing, it's very hard to argue that you're prejudicing your client's interests simply by engaging in the process. –(Robert Baker, attorney)

Similarly, Frank Goldman, an attorney at a large plaintiff's law firm and whom we expected would be

---

[62] Jurisdictions are split on whether, and under what circumstances, parties may be required to disclose seed sets used for model training. See Shannon H. Kitzer, *Garbage in, Garbage out: Is Seed Set Disclosure a Necessary Check on Technology-Assisted Review and Should Courts Require Disclosure Notes*, 2018 U. ILL. J.L. TECH. & POL'Y 197 (2018). Proponents of continuous active learning (CAL), or Tar 2.0 described earlier in the paper, may point to the seed-set disclosure issue as a reason to use TAR 2.0, as it does not require an initial set of documents for training.

more distrustful of the other side than his corporate defense lawyer counterparts, instead reinforced our finding on this issue:

> Q: I guess what I'm getting at is do you think it would be useful for there to be more clear guidance or standards about understanding the other side's TAR process and TAR system?
>
> A: Yes, absolutely. More transparency I think, as a general rule, is better. I think that – I'm pausing because it's a heavy question. What should be happening is your document requests should be honestly and properly and carefully followed and answered. On some level, I care how the materials are gathered and I care what the search protocols are and I want to be really, really strategic and smart, and even if not paranoid, then very, very deliberate and careful. But if I were adjusting a dial, it wouldn't necessarily be so that I could peer into the TAR process of my adversary. It would be so that I could effectively trust and verify their discovery production. There are, I think, other places in the discovery process where that is more real and more inflammatory and more core than peering over the TAR process. (Frank Goldman, attorney at large plaintiff's firm)

CONCLUSION

In conclusion, we highlight two particular issues that we hope will spur further research and action from the legal profession. First, the introduction and increasing popularity of predictive coding systems is reshaping the legal profession in significant ways. Far from being just a new tool in "normal practice," what we have found is that predictive coding—and, perhaps, the broader set of complex, machine-learning-based legal technologies entering the profession—has brought new entities and technical experts into the legal-services ecosystem who are mediating the relationship between lawyers and clients. This new technical tool, combined with cost pressures and ever-increasing volumes of digital data, has created fertile ground for new services and new markets. Thus, these systems call to mind old questions, such as those about contract attorneys and outsourcing of legal work. But these questions get raised in slightly new forms and involve new parties. This has resulted in a reconfiguration of social relations and created new power dynamics, specifically a) new kinds of professionals who have the training and expertise to build and use the tools in ways few lawyers do, and b) a new disciplinary force that corporate clients are using to contain costs.

Second, our discussion of ethical obligations, particularly the duty of technological competence, reveals that more work needs to be done to address potential blind spots at the intersection of professional governance (via rules of professional conduct) and legal decision support technologies. Our interviews found that lawyers are reliant not only on "black box" technical tools but also on other experts. Just in our case of predictive coding for e-discovery, lawyers relied on non-lawyer support staff and vendor judgment for a variety of tasks: system selection (reliant on vendors and, for some, on in-house litigation support staff for early testing), configuration (reliant on in-house technical experts or vendors), and model testing and evaluation (we found no real standards or benchmarking). The exception was lawyer evaluation of system outputs in the case of predicted positive outcomes (but not predicted negatives), wherein lawyers use their own human review to identify false positives of the sort that risk privileged documents being turned over. (Of course, this creates a potentially more problematic dynamic in some instances because it would seem to favor less disclosure by design.)

More importantly, the tools that those non-lawyer experts are plying (under the supervision of lawyers) should be designed in ways that align with lawyers' goals. If we assume that technical experts and algorithmic systems designed to assist lawyers in making legal decisions will continue to increase in volume and sophistication, it is imperative that they account for the professional values and ethical duties implicated by their use. Lawyers need more education and training to better understand machine-learning-based decision support systems. Engineers and developers should also design these systems so they relate to, or at least present in a way that relates to, professionally relevant decision making. For example, while our lawyer respondents were cognizant of traditional privacy and security issues with respect to data storage on cloud-based vendor platforms, they did not seem to be aware of other privacy implications—for example, the risk of model inversion attacks that could allow access to information about training data.