LEGAL E-DISCOVERY SYSTEM

CASE STUDY: 8

The development of artificial intelligence (AI) systems and their deployment in society gives rise to ethical dilemmas and hard questions. This is one of a set of fictional case studies that are designed to elucidate and prompt discussion about issues in the intersection of AI and Ethics. As educational materials, the case studies were developed out of an interdisciplinary workshop series at Princeton University that began in 2017-18. They are the product of a research collaboration between the University Center for Human Values (UCHV) and the Center for Information Technology Policy (CITP) at Princeton. For more information, see **http://www.aiethics.princeton.edu**

PRINCETON

DIALOGUES ON
AI AND ETHICS

New York has the most lawyers per capita of any state in the US. The nearly 180,000 lawyers in New York State must not only compete to represent clients, but must do so while following a strict code of professional ethics. In keeping with the American Bar Association's (ABA) *Model Rules of Professional Conduct*, the New York State Bar Association (NYSBA) adopted a professional code of ethics in 2009 by which all attorneys barred (i.e., licensed) in that state must abide in order to remain in good standing. A lawyer found to be in violation of these rules can suffer penalties ranging from sanctions to disbarment.

NYSBA's Rule 1.1: Competence, requires that lawyers possess the necessary "legal knowledge, skill, thoroughness and preparation."[1] In other words, lawyers must not only be capable, but they must be prepared, their cases thoroughly researched. For litigators, this in part means performing due-diligence in the "discovery phase" of a trial. Litigators are expected to find and assess all materials and information relevant to a client's case through the tedious process of document—or "doc"—review (i.e., fact-finding). When opposing counsel submits a "discovery request," or a demand for specific information bearing on the case, they must then share those findings. NYSBA's Rule 3.4: Fairness of Opposing Party and Counsel stipulates that lawyers can neither "suppress any evidence… [nor] conceal or knowingly fail to disclose" that which they are legally obliged to produce or reveal.[2] All discovery that is deemed relevant to the opposing counsel's request and is not protected by attorney-client privilege, must be shared.

At big law firms, doc review has traditionally been conducted by first- and second-year associates, who spend many of their working hours poring over data. When opposing counsel makes a discovery request, those same attorneys must evaluate each document to determine, not only whether it is harmful or helpful to their client, but also whether it is responsive to the request (or not) and/or privileged (or not). Given time constraints and the quantity of information to be reviewed in even a moderately-complex case, junior associates may have mere seconds to scan each document and make such evaluations. Where the answers are not straightforward, lawyers may err in the direction of nondisclosure rather than risk sharing irrelevant information or that which is protected by privilege. Doc review is thus an imperfect, arduous process that exacts a heavy toll from the junior associates responsible for this work.

Advances in AI have begun to change the nature of doc review with the introduction of electronic discovery, or e-discovery, systems. Commonly known as technology-assisted review (TAR), these systems apply supervised machine learning (ML) or rule-based techniques to search electronic data for information pertaining to particular civil or criminal proceedings. TAR also uses textual analysis to evaluate data in terms of its likelihood to be privileged and its responsiveness to a discovery request

Discussion Question #1:

Burnout is common among junior litigators enmeshed in doc review. The work can feel dehumanizing. And because expensive human lawyers tend to perform poorly at doc review in terms of both false negatives (i.e., identifying material as irrelevant and/or privileged when it is not) and false positives (i.e., identifying material as relevant and/or non-privileged when it is not), many firms now outsource this work either to overseas contractors or AI systems. What are the advantages and disadvantages of these two types of outsourcing? Is there a benefit to overseas outsourcing over TAR, or vice versa?

[1] https://www.nysba.org/WorkArea/DownloadAsset.aspx?id=50671, pp. 11.
[2] https://www.nysba.org/WorkArea/DownloadAsset.aspx?id=50671, pp. 117.

Discussion Question #2:

Lawyers must turn over information that meets the standards of a discovery request. But in borderline cases, humans use their judgment to assess data (often erring on their clients' side). TAR limits the use of such discretion. Is that a good or bad thing?

TAR has grown popular in the legal market for several reasons. For one thing, cost-conscious clients who balk at the $300-$400/hour rates that big law firms charge for junior associates to conduct doc review often request cheaper alternatives. Law firms, too, can appreciate the savings of switching to AI. Furthermore, firms may feel pressured to adopt these technologies in light of their professional ethical obligations. NYSBA's Rule 1.1 states that lawyers must not only *be* competent, but they must *remain* competent in a changing world. This means staying up-to-date regarding changes in the law relevant to their practices, attending continuing legal studies courses *and* modifying their methods in accordance with the evolving technological landscape.

Rule 1.1, Se. 8 insists lawyers "keep abreast of the benefits and risks associated with technology the lawyer uses to provide services to clients or to store or transmit confidential information…"[3] Lawyers are thus expected to evaluate new technologies that may improve the discovery process. Where they are deemed actual improvements, those technologies should be adopted. While the courts haven't yet forced widespread implementation of TAR, recent decisions suggest that such a shift may be on the horizon. In 2012, *Da Silva Moore v. Publicis Groupe* ruled that, as TAR "appears to be better than the available alternatives," it "should be used in appropriate cases."[4] Only four years later, *Hyles v. New York City* concluded that, for most cases today," TAR "is the best and most efficient search tool."[5]

Discussion Question #3:

New York-barred lawyers have a professional obligation to use new technologies that aid in client representation. But as lawyers are rarely technologists, they may find it difficult to sort through the available options and determine which best fits their needs and values. Instead, they may choose TAR systems based on reputation, cost or even personal connections to vendors. This is no different from how many non-experts decide between AI systems. Given the lack of technical expertise in various professions that use AI, how can/should people make decisions between different systems? To what extent should they be responsible for the outputs of systems they don't fully understand?

Discussion Question #4:

Law practices come in all shapes and sizes, from cash-strapped to wealthy, solo practitioners to global firms employing thousands. Should TAR become required in the legal profession, it will impact these organizations differently. What are some advantages and disadvantages of using TAR relative to an organization's size and resources?

Johns, Morton and More, a prestigious New York City-based law firm recently adopted an expensive, top-ranked TAR, YouTopia. Some of the more senior attorneys balked at the idea of outsourcing doc review to AI, with several lamenting how easy today's lawyers now have it. The vast majority of the firm embraced the technology, however, noting that it would save a great deal of time and money in the long run. The most junior associates, especially, were excited about the possibility of a world with only limited doc review. While they would remain "in the loop," reviewing and spot-checking YouTopia's findings, these attorneys expected to finally have time to train in the more interesting aspects of the law.

[3] https://www.nysba.org/WorkArea/DownloadAsset.aspx?id=50671, pp. 13.
[4] See *Da Silva Moore v. Publicis Groupe*, No. 11 Civ. 1279 (ACL) (AJP) (2012).
[5] *Hyles v. New York City*, No. 10 Civ. 3119 (AT) (AJP) (2016).

Shortly after the firm's purchase of the YouTopia system, it was asked to handle a major case. Six female executives from Tudor Industries came to Johns, Morton and More, wishing to file charges against their previous employer and its CEO, Hank Ocho, for sexual harassment and the proliferation of a hostile work environment (despite repeated complaints). Having been unable to bring about criminal proceedings, they were interested in initiating a civil suit through which they could at least receive damages for past harms.

The case was assigned to Tommy More, a senior associate at the firm and the son of one of the partners, Agnes More. The younger More was considered a rising star in the legal community, not just because of his keen legal mind, but also because of his reputation for ethical dealings and a strict adherence to the rule of law. He immediately set his team to the task of finding everything that could possibly impact the suit against Tudor Industries and its CEO.

As Johns, Morton and More's e-discovery group was still in the process of getting YouTopia up-and-running, the team began their internal doc review the old-fashioned way, conducted by humans. The junior attorneys sorted through everything the women had in the way of paper trails—both digital and analog—including personal emails, text messages and phone records. There was no smoking gun, but they did uncover several emails with sexual innuendo that seemed to support the women's accounts. The more senior associate who reviewed these findings was satisfied that this case should be an easy win.

Tim Cromwell, lead counsel for Ocho and Tudor Industries, shortly sent over some interrogatories—i.e., specific questions about the case—and submitted a discovery request to Johns, Morton and More. The request demanded: 1) all email and text exchanges between the six women named in the suit and Ocho, Tudor Industry's Board of Directors and the Human Resources Department between the dates of each woman's first day of employment and her last; and 2) all email and text exchanges during those time periods that included Ocho's name and certain specified sexual terminology. The request also specified that More's team use the YouTopia system it had recently purchased to conduct the doc review, as it was known for its high level of accuracy compared to both humans and other TAR systems on the market.

Confident in his client's claims and his team's doc review skills, More was happy to oblige. He was, however, much less happy when he received the results. YouTopia uncovered text exchanges between two of the women, which the lawyers had missed, suggesting they may have coordinated the details of their allegations. And in analyzing the metadata, YouTopia was able to place several women outside of the office at times they claimed in-person harassment to have occurred, thus calling into question the overall accuracy of their accounts of what had transpired. Interestingly, YouTopia *did not* identify the damning messages that the lawyers had found, possibly suggesting the limitations of relying on algorithms in such contextually-dependent areas of the law as harassment.

Discussion Question #5:

"Accuracy" is not always clear-cut in AI. YouTopia successfully identified certain kinds of data that human lawyers could miss (e.g. metadata) but performed worse with other kinds (e.g. sexual innuendo). Given this incongruity, what does it mean to talk about accuracy in terms of AI? Is there a better alternative for measuring a job well done?

It is a lawyer's fiduciary duty to do what's best for the client, but lawyers must also turn over all responsive, non-privileged discovery. In this case, the information YouTopia found might be all the defense needed to cast doubt on the veracity of the women's allegations. In the ambiguous area of "he said, she said" sexual harassment law, this added doubt might mean the difference between the accusers being summarily dismissed or not, even if the revealed discrepancies did not fully undermine their claims. Many of the attorneys felt torn between their conflicting obligations to the clients and opposing counsel. Several lamented the adoption of YouTopia in the first place; had it not been available, the inconvenient evidence may never have been unearthed and there would have been no conflict. These attorneys voiced their concerns to More and wondered how to mitigate such issues in the future.

Ethical Objection #1: Accuracy V. Ambiguity

More's group would never argue against the value of accuracy and inclusivity in doc review. Litigators must learn all they can to avoid surprise evidence coming out at trial. However, lawyers may still want a little "wiggle room" regarding the facts of the case. While it is the judicial system's aim to find truth and promote justice, in an adversarial legal system—such as exists in the US—it is the lawyer's responsibility to provide the client with the strongest possible defense. Lawyers sometimes interpret this duty to demand not knowing *everything*—or not knowing with absolute certainty—given that they must share all relevant, non-privileged discovery requested by opposing counsel. Similarly, lawyers may desire flexibility to determine what evidence is or is not relevant and/or privileged, especially for borderline cases. The YouTopia skeptics claimed that highly-effective TAR compromises attorneys' ability to fulfill their professional obligations by leaving little room for ambiguity where it may benefit the clisent, both in fact-finding and also in assessing information for privilege and relevance.

Ethical Objection #2: How Good Is Too Good?

For many of the lawyers on More's team, the issue with YouTopia wasn't that it performed poorly at doc review—although it did for certain requests—but that it performed *too well*. YouTopia was not only better than most other TAR systems, but performed faster and with a lower rate of error than human lawyers. While such high accuracy may initially appear to be an obvious advantage, to the extent that knowing too much could harm clients' interests, the lawyers were not so sure. After all, the NYSBA's professional obligations regarding discovery requests had been designed with human lawyers and their capabilities in mind. They wondered whether TAR that significantly outperforms humans should be held to different standards. Or perhaps, they suggested, TAR with superhuman abilities should be abandoned altogether. After all, most firms that had adopted TAR systems had opted for cheaper models that performed more similarly to humans. In that case, was there really an advantage to using the more expensive, effective YouTopia system? Indeed, clients might prefer a cheaper model if given the choice.

Ethical Objection #3: Displacement Of Workers

With concerns about how YouTopia interacted with lawyers' professional obligations dominating discussion threads at the firm, an additional consideration regarding TAR began popping up. While many of the lawyers at Johns, Morton and More were initially pleased with the cost and labor savings that YouTopia afforded, the early boasts about wanting to delegate doc review to "robots" had died down. As junior associates were no longer needed to conduct doc review, they were free to perform the kind of substantive work that they were eager to begin. Unfortunately, it quickly became clear that there was not enough of that kind of "good" work to go around. Johns, Morton and More were not looking to let anyone go at this point, but the managing partners were considering whether they ought to offer fewer positions in the next hiring cycle. If all firms behaved similarly, some wondered what long-term effects this would have on a profession in which the supply of lawyers already outstrips the demand by firms. More immediately, what would happen to all the new lawyers with nowhere to go and mountains of student debt that had been acquired under the expectations of there being jobs available for them post-law school? Who was going to take responsibility for them? What responsibilities, if any, do the individual firms whose adoption of TAR has displaced them have to the upcoming generation of lawyers?

More was disconcerted by his team's complaints, not because they were unjust, but because they highlighted a challenging ethical tension in his work. His professional duty was to his clients, yes, but he could not wrap his head around the idea that knowing more was not inherently a good thing. Indeed, isn't knowing as much as possible also part of his professional obligation? Is it not a moral imperative as well? YouTopia had changed what "as much as possible" meant in the legal scenario, but More did not believe it altered the underlying premise that truth-seeking is inherently valuable. More wished to serve both his clients and the "truth, the whole truth and nothing but the truth," but what if the two were incompatible?

In his turmoil, More sought guidance from a former law school friend now employed at a competing firm, Liz Barton, to ask if she had ever encountered this conundrum. She replied that she had not, but Barton also indicated that her firm used a different TAR, one with accuracy ratings more on par with human capabilities. Indeed, according to a recent study by the US Department of Commerce's National Institute of Standards and Technology (NIST), most TAR products on the market show low accuracy, meaning that they underperform compared to humans. Given that this was the industry standard, Barton suggested that perhaps Johns, Morton and More adopt a more "human" AI system. She had become friendly with the brand rep for her firm's TAR, who frequently bought the lawyers dinners and drinks while selling his wares, and Barton said she would be happy to make the introductions.

Discussion Question #6:

YouTopia could be said to have been "too good" at its job, which arguably undermined the interests of those it was meant to serve (i.e., More's clients). Similar charges have been leveled against facial recognition systems used in police surveillance. To the extent that they are "too good," they may enable practices that harm the very communities they are developed to serve. Thinking of YouTopia, are there other possible harms associated with being "too good" that haven't been identified here?

Discussion Question #7:

Some TAR systems outperform humans and some are on par with human doc reviewers, but many perform considerably worse. Performance metrics often correlate with cost. How should we think about the adoption if this technology given the range of accuracy and costs, and the lack of transparency about both? Should clients be informed about both accuracy and cost of the TAR systems used for their cases?

More was deeply uncomfortable with his friend's suggestion that Johns, Morton and More switch to a less effective TAR, but knew that several of his colleagues had had the same thought. Indeed, the e-discovery group was already scouring the NIST report for a TAR towards the lower end of the accuracy scale. And while it was too late to switch away from YouTopia in the case against Ocho and Tudor Industries, if the firm adopted a cheaper, less accurate TAR for future cases, they could enjoy even greater cost savings without the concerns associated with knowing too much. (At the same time, they might also be able to put off the question of how widespread adoption would impact the profession in the future)

More disagreed and defended YouTopia. The truth may sometimes hurt, he argued, but that does not justify lawyers closing their eyes to it. Unfortunately for More, Johns, Morton and More was ultimately unpersuaded and decided to abandon YouTopia and adopt a less capable TAR.

There was nothing More felt he could do in good conscience at that point but resign his position, and he considered returning to school to study moral philosophy. Before leaving his office, More wrote one final email to the firm. "I leave as the good servant of my clients, whose interests I have always had at heart," he wrote, "but my principles must come first." His former colleagues felt he might have been overdoing it.

Reflection & Discussion Questions

Accuracy: Accuracy is often treated as a non-controversial goal of AI development. But increases in accuracy may not be such an unambiguous good. Facial recognition systems, for example, have vastly improved in terms of accuracy in recent years. But those improvements have sparked debate about whether greater accuracy is needed or advantageous, especially in light of how these tools might be used against vulnerable populations. Wasn't the "good enough" version of yesteryear sufficient for the basic tasks facial recognition might perform? What kind of functions are made possible only by more accurate facial recognition systems? Are these functions necessary? Might they cause harm? In the case of YouTopia, the system's high accuracy (regarding certain tasks), as compared to both human lawyers and other TAR, may have been a double-edged sword. To the extent that superhuman accuracy risked harming the clients' interests, some argued that they would have preferred a less advanced TAR.

- Are there circumstances in which a high-performing AI system ought to do less than its best? Are there reasons why engineers might wish to develop products that perform less well than the technology allows? Consider both financial and non-financial reasons in your response.
- Despite its high marks for accuracy across several dimensions, recall that YouTopia struggled to identify sexual innuendo that human lawyers had previously found. Does this suggest a need for human-machine collaboration? Is there a way to narrow the accuracy gap for YouTopia?

Competence: Competence is listed as Rule 1.1 by both the ABA and the NYSBA, underlining its special importance in the role of being a lawyer in the United States. However, given their training, most lawyers are not also technology experts and may, therefore, lack the technological competence to make informed decisions about the adoption and use of TAR. This problem extends to other forms of AI now commonly employed in the legal system as well (e.g. risk assessment tools that judges use to determine bail applications). Such AI can have profound impacts on those to whom it is applied, as well as society at large. And yet the legal professionals using them may not fully understand how they work, which values they promote or the costs/benefits associated with particular models.

- A machine learning (ML) system is only as good as the lawyer (or judge) operating it. If we accept that most legal professionals lack some technological competence, how should the decision to adopt a particular AI system be made in this field? What role should the values of transparency and accountability play? Does your answer differ whether you're considering a private law firm, a public defender, a local court, an appellate court, etc.?
- One option to address the disconnect between the use of AI in the legal system and the lack of technological competence among legal professionals is to change the training requirements for the latter. If you were to design an educational initiative to inform legal professionals about the AI in their field, what would it look like? Where would you begin?

Job Displacement: When considering how AI may contribute to the displacement of workers, litigators might not be the first group that automatically comes to mind. Many other industries—including manufacturing, long-haul trucking, etc.—may face both practical and existential crises even sooner and more critically, as AI begins to play an increasingly prominent role in their fields. But while it may take longer for the economic effects of AI to be felt among lawyers and other professionals, it is important to keep in mind that no industry is entirely immune. To the extent that the introduction of AI to the workplace is bound to disrupt the current balance of labor and displace workers, it is a societal issue, and not one confined to any particular industry.

- The lawyers at Johns, Morton and More raised some important questions about both the long- and short-term impacts of introducing an AI system to their workplace. They recognized that in both scenarios, there are bound to be lawyers who lose out because of TAR. Critically, they tied this back to the question of responsibility: Who or what will take responsibility for those who are disadvantaged by technological advancements? They suggested that the firms, themselves, might bear some responsibility. Do you agree or disagree? Why?

- AI optimists may argue that job displacement is not inherently a bad thing, as long as workers are given the opportunities and necessary training for them to pivot to other sectors of the economy. Does this seem like a satisfactory response? Would debt-ridden law students be justified in rejecting such an approach?

AI Ethics Themes:

Accuracy

Competence

Job displacement