Intersect, Vol 15, No 1 (2021)

Accessing Justice Through Technology: An Interview with Professor David Engstrom

Katie Yoon
Stanford University

David Freeman Engstrom is a nationally recognized expert and award-winning scholar in civil procedure, administrative law, and constitutional law. His current work focuses on the intersection of law and artificial intelligence. He is working on a project on the effects of continuing advances in "legal tech" on the civil justice system and the governance, lawyering, and access to justice challenges posed by AI. As part of that work, he is serving on the State Bar of California's Closing the Justice Gap Working Group, tasked with proposing reforms to foster innovative legal service delivery systems. During 2018-2020, he served as a principal advisor to the Administrative Conference of the United States on the project, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, which garnered national media attention and remains the most comprehensive treatment of the subject to date.

At Stanford, Professor Engstrom co-directs the Stanford Center on the Legal Profession. From 2018 to 2021, he served as an Associate Dean at Stanford Law School and led an initiative charting the school's future work around digital technology. He is a faculty affiliate at CodeX: The Stanford Center for Legal Informatics, the Regulation, Evaluation, and Governance Lab (RegLab), and the Stanford Institute for Human-Centered AI, where he also chairs the Technology Policy Governance committee.

Beyond teaching and research, Engstrom has served as counsel or consultant to a wide range of public and private entities and is a frequent amicus before the U.S. Supreme Court. He is an elected member of the American Law Institute and a faculty affiliate at the Stanford Human-Centered AI Initiative and at CodeX: The Stanford Center for Legal Informatics. He holds a J.D. from Stanford Law School, an M.Sc. from Oxford University, and a Ph.D. from Yale University.

Yoon: You have a truly impressive academic and professional career. From my research, I see that you initially studied history and political science. How did you first grow interested in technology and its intersections with the law?

Engstrom: I initially became interested in law and technology when I learned that the federal government was developing tools that use machine learning to perform some really important governance tasks, such as adjudicating social welfare benefits or choosing regulatory enforcement targets. And it just became clear that the technology was sweeping into the government in consequential ways. So, I started to study it. I have a PhD in political science, so I already have quantitative training. And that made it possible to learn machine learning and to see how it is and isn't different from the standard inferential statistics that I learned. I've learned enough machine learning to understand some of the potential and some of the perils of these tools.

My initial research led to a big project here at Stanford, generously funded by HAI, that involved some of my terrific colleagues including Dan Ho at the law school. Also, Tino Cuéllar (Mariano-Florentino Cuéllar), who's a former professor, and until recently was a Supreme Court justice in California. He just stepped down from that role to take a position as head of the Carnegie Endowment for International Peace. And then also a colleague from NYU named Kathy Sharkey (Catherine Sharkey). So, we jumped in and did the first big comprehensive study of government use of these tools.

From there, it wasn't hard to take all this knowledge that I had gained on the use of AI-based tools by government agencies and start thinking about how they will impact another part of the public sector, the justice system. How will they impact what courts do? How will they impact how lawyers do their jobs? Or how might they be used to assist self-represented litigants? Self-represented litigants are pro se litigants—litigants who don't have access to a lawyer, who can't afford a lawyer, who can't find a lawyer. How might these tools help them and close what many see as a significant access to justice gap in the United States? And so that's where a lot of my more recent work has taken me.

Yoon: Can you perhaps expand on that? How did the idea for your policy practicum initially come about?

Engstrom: The policy lab I'm currently co-leading here at the law school involves students from all over the university. So, it's not just law students—it also includes undergraduate and graduate students across disciplines. But the idea for that came from my interest in how technology was starting to pervade the justice system. There's this very significant access to justice problem in the United States. I can try to convey it with a single statistic: in the roughly 20 million civil cases filed in state and local

courts per year, in roughly three quarters of them one of the parties—and in some cases, both of the parties—doesn't have a lawyer.

We have this very stylized, even mythical model of what the American legal system looks like. And we think of great legal dramas, like *A Few Good Men* where there's a courtroom trial between two highly skilled lawyers who are essentially duking it out in front of a judge. But that just doesn't describe the overwhelming majority of cases within the American legal system, especially at that state and local level. And these cases are some significant cases. They tend to be low stakes viewed from the outside, which is to say, there may not be that much money at issue. These aren't billion-dollar cases. But they're very consequential and very important for the parties themselves.

So, what are these cases? The majority of them tend to be some combination of the following. First and foremost, at least in numerical terms, are collections cases—for example, a debt collector going after a debtor for credit card debt or for failure to make a car payment or a mortgage. They also include eviction cases where a landlord is trying to evict a tenant. And they include a lot of family law cases—for instance, where an individual who is suffering abuse at the hands of a partner is trying to get a protective order, to get protection from that person. And so, the gap is that there are all these people with these very significant justice needs. They tend to be relatively low value cases, at least in monetary terms, and so it simply doesn't work to pay a lawyer. There's no plausible business model whereby a lawyer can provide representation to a lot of these people. If you're sick, if you're facing a debt collector, that probably means you don't have any money in the first place, and so therefore you can't pay for a lawyer.

The impetus for the policy lab was, “How do we move the dial on this?” There have been lots of past efforts to try to close the justice gap. Those efforts have included trying to find lawyers who are willing to volunteer their time or finding ways to pay for lawyers. And they have included allowing non-lawyers to provide representation. You can think of nurse practitioner types who have specialized expertise in a particular legal area, like collections or evictions, and who might be able to provide legal services to people who would be well served by someone who isn't a full-fledged lawyer but could still help them navigate the system.

But a lot of people think we can't lawyer our way out of this crisis. And, in fact, we can't even *human* our way out of this crisis. The needs, and the numbers, are just too great. There's a view that what we need to do is find good high-quality software applications that can assist self-represented litigants in their efforts to navigate the system on their own.

And so that's what our policy lab is about. There have been lots of efforts to build software applications that try to help a self-represented litigant to complete a form, or write a complaint or an answer, or any of the other types of pleading or legal documents they’ll need to vindicate their rights in court. There have been lots of efforts to do it. But they've

largely failed because they've been limited to a single jurisdiction. They've tended to be intra-jurisdictional, small-scale efforts, and, as a result, they couldn't attract a high-quality software application that could actually be of use to a litigant who knows very little about the law.

So, the idea for the policy lab came about with my colleagues Mark Chandler and Margaret Hagan—both co-leaders of the project. Chandler is a longtime Silicon Valley General Counsel who stepped down as the top lawyer at Cisco, which of course, is an enormous and powerful tech company. So, Mark was, and is, a very skilled and very effective lawyer. Margaret Hagen is the head of the Legal Design Lab at the d.school and the law school here at Stanford and one of the leading thinkers on how to redesign our courts and legal system. We started by asking why these various past projects have failed, and what we could do to try to mount a more successful version. And what we decided was that maybe the problem is a lack of scale. Maybe the problem is that, because the American system of federalism has all these states and all these different jurisdictions with different laws and different filing standards and different data infrastructures, there's never been sufficient scale to attract tech providers and convince them to invest their time and energy into a high-quality piece of software.

This project is thus all about trying to lead with scale. It's about trying to identify a set of states, and one or more types of claims, that are ripe for automation. And then we can use Stanford's convening power to bring those jurisdictions together and get them to commit to standardize and to hold constant, whatever their laws are, and whatever the different procedures are. Then we need to convince them to bring into alignment their data infrastructure, even something as simple as the data fields that would be used to populate the forms that self-represented litigants need to file. And then working with a tech company, we could help to develop a high-quality piece of software that would be easy to use, that would be usable across multiple jurisdictions, and therefore could help a lot of self-represented litigants who can't get counsel to navigate the system.

Yoon: Thank you for that very detailed explanation. So, you mentioned how you've only recently got interested in tech. But even before then, how have you seen the legal tech industry grow over time since your start as a lawyer and litigant?

Engstrom: We've most certainly seen increasing use of technology by lawyers to do their work. We sometimes refer to those tools using a catch-all phrase "legal tech." Legal tech, by the way, could mean tools that help self-represented litigants. But more commonly when we say legal tech, we mean *lawyer-driven legal tech*, which is to say a tool that either augments what lawyers do or maybe even substitutes for what lawyers do but is still operated by them on behalf of clients. Going forward, these lawyer-

centered tools are going to be a more and more pervasive part of legal practice and the legal system.

To this point, the main area where we've seen legal tech tools really penetrate the system is eDiscovery, sometimes called technology-assisted review, or TAR, and sometimes called predictive coding. Whatever the label, we're talking about the use of a machine learning tool to flag documents for relevance and privilege. In a litigation, there are often millions of documents in hard drives in a defendant's system, and the plaintiff serves requests on the defendant for any of those documents that relate to the litigation. And the defendant has to meticulously go through those millions of documents and figure out which ones are and aren't relevant to the case and which ones are and aren't *privileged*—that is, aren't cloaked by various rules that allow a defendant to keep certain documents secret, usually because they involve attorney-client communications.

These tools are powerful, because it means lawyers need only label a subset of those documents. So, if there are 10 million documents that are candidate documents, lawyers could label 1000 of them, or 5000 of them. You need a relatively small team of lawyers working for a few days to put those labels on those documents, not the months it would take to review all of them. Then you can use a supervised machine learning model to label the rest, to flag the remaining 10 million for relevance or privilege. Now, as you and your readers probably know, there's lots of tuning that must be done to that ML model. This is not just a turn-the-crank sort of thing. But the point is that this is a potentially powerful tool, and it can cut costs and litigation massively.

The view is that these TAR systems, so long as implemented carefully, are as accurate or even more accurate than manual, human, eyes-on review. And yet, they operate at a fraction of the cost. And that can really revolutionize the litigation system in lots of different ways. So much of our thinking about civil procedure in recent decades has focused on concerns about litigation costs. Litigation costs have not only been a real fixture of debate in American politics, but also in the evolution of legal doctrine over the last 40 or 50 years. And so a new technology that can dramatically reduce the cost of litigation would be important, an absolutely transformative technology. So that's a place where sophisticated new technologies are already in use in the legal system, and this use case will only grow in importance in the coming years.

But as you and your readers know, natural language processing has gotten more and more potent, and more and more powerful. Think here of Google BERT or the GPT-3 foundational models. As NLP has gotten more and more powerful, now people are thinking about the possibility of other applications within the legal system. After all, the legal system is a system that trades in words. That's what legal systems are. And so, NLP could automate lots of higher order legal cognitions.

So, you can imagine an NLP tool that can do something we would call *legal analysis*—deciding which cases are most relevant to the instant case or which cases would allow a lawyer to win the case or fend off the arguments of the other side. We could also imagine a tool using NLP that could provide a *case outcome prediction*—a prediction as to how the case would come out if it was litigated to a judgment in court. And if you think about that, that's the essence of what lawyers do. One of the most important types of counsel that a lawyer provides to a client is: What are my odds? What are my chances of winning if I go forward with this? Or what are my chances of losing if I go forward with this? And how much might I lose, or how much might I win? These are higher-order legal cognitions that, for the moment, only lawyers can do. But we're already starting to see NLP tools that can make some progress on these sorts of tasks, especially in relatively narrow technocratic areas of the law, like tax law or housing law.

Yoon: Wow, it's super incredible to see you can even predict outcomes using AI! I assume you've probably faced challenges persuading a lot of people in the legal space to become more receptive to such technologies, especially with fears of AI replacing jobs. Do you have general advice for bridging divides across these different disciplines? And have you tried to overcome barriers in terms of reticence towards new technologies and AI?

Engstrom: I talk to lawyers a lot and I talk to judges a lot about these things. So that's a kind of bridging, I suppose. I've also tried to think about how we can do more to bring together technologists and lawyers, including through classes here at Stanford. I think that's a second critically important way we need to do a better job of bridging within the law and technology space.

Both kinds of bridging are important, because lawyers and judges might just be culturally predisposed against technological innovation that might make the system better. We lawyers tend to be Burkeans. We tend to be sort of small-c conservative in our thinking. We're trained to see around corners and to spot reasons not to do things. And so, there's a real cultural problem within the legal system, which is at least part of the reason why lawyers, but especially maybe judges, are averse to new technologies.

In some cases, of course, skepticism is healthy. There are all sorts of interesting reform moves right now within the courts that may or may not be a good thing.

A good example is increasing experimentation by courts with ODR, or online dispute resolution. To this point, most online dispute resolution, at least in the court context, was nothing more than an asynchronous gathering place where the disputants can come and try to bargain their way to a settlement. We could call that ODR 1.0. But there's this new kind of ODR that's more and more possible as a result of advances in NLP and

other techniques that I call ODR 2.0. And that's a form of online dispute resolution that has an algorithmic tool built on top of it. That might be a blind bidding system, or it might be some other way to try to find overlap in the party's views of a case and try to help them get to settlement.

But going forward, it seems like it will be possible to arm disputants in ODR systems with a prediction as to how their case might come out. This would be a very potent form of ODR. It could help to solve the access crisis that we've already talked about here today. Remember that one of the most important things lawyers do is provide counsel on the likely outcome of a case. And so, there's lots of promise built into ODR.

But there are also lots of people who rightly say: How are we going to govern this? How are we going to decide what the algorithmic system pushes to the disputants by way of information? How can we ensure that it's a fully validated system? How can we be sure that it's fair and equitable and doesn't perpetuate different forms of bias that might exist in the data?

These are important questions because these ODR systems are very likely to be deployed in areas like debt collections and evictions, where the sheer number of cases is overwhelming the system. The question becomes, should we worry that ODR will just become an efficient assembly line for debt collectors to get their judgments against debtors and garnish your wages once they've gotten a judgment? Or a super-efficient assembly line for landlords to achieve evictions against tenants? And so, we need to think hard about that. If we think this is going to solve an access crisis then maybe, we should be putting lots of energy into thinking about how we can build it out and make it more and more pervasive. But if instead, what we're doing is simply increasing the efficiency of a system that some might see as already unfair, then maybe we should think twice about putting lots of our energy and industry towards it.

Yoon: Another question that arises is, when discussing these sorts of legal issues and technological applications, how do you account for the potential that they might exacerbate such pre-existing social and racial inequities? Is there a systematic way in which to consider all these different social issues?

Engstrom: I think that's an important question. And I think it relates nicely to what I just said about ODR. If we think that system is already one in which there exists a significant degree of bias against litigants, especially persons of color within the system, then a system that leverages past decisions to try to give future disputants a prediction as to how their case comes out is just going to perpetuate that bias.

There's also a further social context here that's worth mentioning. There are plenty of people who say that the government has done less and less over time for a variety of reasons to insulate individuals from shocks to their income. Put another way, the social welfare state has grown

thinner and thinner in the United States, and in lots of other countries as well. And the result is that we've shifted a lot of risk to individuals and away from government in a way that's quite different from the original vision, say, during the New Deal, when we first built a modern social welfare state. A health problem can suddenly create significant financial problems for individuals. And so maybe ODR systems are not a salutary innovation but, instead a terrible failure of imagination on our parts. Because maybe what we need is a social welfare system that better insulates people from those shocks to their income. And maybe the point is not to create a more efficient and better judicial system, or a more efficient way for landlords to evict tenants who can't pay, but to rethink how we try to protect people from those income shocks in the first place.

Yoon: Finally, do you have any last words you would like readers to leave from reading this interview?

Engstrom: I will leave with this, which is, I think the American legal system is at a pivotal moment because of the advent of all these new and potent technologies. I think the next 10 or 15 years is going to be an absolutely critical period in the life of American law and in the life of the American legal system. As a result, I think it's an all-hands-on-deck sort of moment. It's going to take lawyers, social scientists, folks who specialize in the humanities, and technologists to navigate our way into that new and digitized world in some sort of responsible way.

And given that it's an all-hands-on-deck moment, I always try to beat the drum for more involvement of non-law students here at the law school. So, I would just point out, we have 20 plus students right now in the policy lab that I described earlier on how we could reform civil filing systems to better serve self-represented litigants. I would say half of those are law students, but the other half are not, and they come from all over Stanford, from the Engineering quad to the Philosophy department. So, I guess what I would say to the readers is, this is an important moment. And there's a lot of really impactful work going on at Stanford on these issues, especially at the Law School. I would invite undergraduates, or graduate students, or anyone who reads this at Stanford to come look at the policy labs, for instance, that we run here at the law school which provide some amazing opportunities for folks with all different skill sets to get involved in thinking about what that digital future of the American legal system is going to look like.