

Oyez! Robot

Author : Michael Madison

Date : January 24, 2020

Richard M. Re & Alicia Solow-Niederman, [Developing Artificially Intelligent Justice](#), 22 **Stan. Tech. L. Rev.** 242 (2019).

How and why does it matter that humans do things that machines might do instead, more quickly, consistently, productively, or economically? When and where should we care that robots might take our jobs, and what, if anything, might we do about that?

It is the law's turn, and the law's time, to face these questions. Richard Re and Alicia Solow-Niederman offer an excellent, pragmatic overview and framework for thinking about artificial intelligence (AI) in the courtroom. What if the judge is a (ro)bot?

The general questions are far from novel, and there is no shortage of recent research. Facing the emergence of computable governance in the workplace and large swaths of social life, for the last twenty years legal scholars, historians, and researchers in science and technology studies have been exploring "algorithmic" decision-making in computer networks, social media platforms, the "gig" economy, and wage labor.

Yet the application of automation to the law feels different, disconcerting, and disruptive for added reasons that are not always easy to identify. Is it the central role that law and legal systems play in constructions of well-ordered modern society? Is it the combination of cognition and affect that define modern lawyering and judging? Is it the narrative of law as craft work that still underpins so much legal education and practice? Those questions form the conceptual underpinnings of Re and Solow-Niederman's work. They are after a framework for organizing thoughts about the answers, rather than the answers themselves.

Organizing a framework means pragmatism rather than high theory. The problem to be addressed is "how can we distinguish the benefits from the harms associated with automated judging?" rather than "What defines the humanity of the law?" Re and Solow-Niederman address courtroom practice and judicial decision-making as their central example.

The article proceeds elegantly in a handful of steps.

First, Re and Solow-Niederman propose a reconfigured model of systems-level interactions between "law" and "technology," shifting from the law as a set of institutions that "responds" to technological innovation (a linear model, labelled "Rule Updating") and toward law as a set of institutions whose capacities co-evolve with technological innovation (a feedback-driven model, labelled "Value Updating").

Within the Value Updating model, the article addresses adjudication, distinguishing between stylized "equitable justice" and stylized "codified justice." The former is usually associated with individualized proceedings in which judges apply legal rules and standards within recognized discretionary boundaries. The latter is usually associated with the routinized application of standardized procedures to a set of facts. The justice achieved by a system of adjudication represents a blend of interests in making accurate decisions and making just decisions.

Re's and Solow-Niederman's concerns arise with the alignment and reinforcement of codified justice by algorithmic systems, the "artificially intelligent justice" of their title. They acknowledge that what they call codified justice is not new; they invoke precedents in the federal sentencing guidelines and matrices for administering disability benefits. Nor

is codified justice, in its emerging AI-supported forms, temporary. Algorithmic judging supported by machine learning is here to stay, particularly in certain parts of criminal justice (for example, parole and sentencing determinations) and benefits administration, and its role is likely to expand.

Re and Solow-Niederman argue that the emergence of AI in adjudication may shift existing balances between equitable justice and codified justice in specific settings, in ways that key into macro shifts in the character of the law and justice. Their Value Updating model renders those shifts explicit. With AI-based adjudication, they argue that we may see more codified justice and less equitable justice. Why? Because, they note, motivations for adoption and application of AI to adjudication are tangible. Codified justice promises to be relatively cheap; equitable justice is relatively expensive. Firms are likely to promise and to persuade, rightly or wrongly, that AI may deliver better, faster, and cheaper decision-making at scale.

The article is careful to note that these shifts are not inevitable but that the risks and associated concerns are real. Perhaps the most fundamental of those concerns is that AI-supported changes to adjudication may shift “both the content of the law and the relationship between experts, laypersons, and the legal system in democratic society” (P. 262) in systematic ways. Decision-making and adjudicative outcomes may be incomprehensible to humans. Data-driven adjudication may limit the production or persuasiveness of certain types of system-level critiques of legal systems, and it may limit the extent to which rules themselves are permitted to evolve. Reducing the role of human judges may lead to system-level demoralization and disillusionment in society as a whole, leading to questions of legitimacy and trust not only with respect to adjudicative systems but regarding the very architecture of democracy. To paraphrase Re’s and Solow-Niederman’s summation: if robots resolve disputes, why should humans bother engaging with civil society, including fundamental concepts of justice and the identity and role of the state?

Re and Solow-Niederman conclude with their most important and most pragmatic contributions, describing a range of stylized responses to AI’s promise of “perfect enforcement of formal rules” (P. 278) that illuminate “a new appreciation of imperfect enforcement.” (*Id.*) Existing institutions and systems might be trusted to muddle through, at least for a while, experimenting with AI-based adjudication in various ways without committing decisively to any one approach. Alternatively, equitable adjudication could be “coded into” algorithmic adjudicators, at least in some contexts or with respect to some issues. A third approach would involve some systematic allocation of adjudicative roles to humans rather than machines, a division of labor approach. A final response would tackle the problems of privately developed robot judges by competing with them, via publicly supported or endorsed systems. If you can’t join them (or don’t want to), beat them, as it were. As with the article as a whole, this survey of options is inspired by broad conceptual topics, but its execution has an importantly pragmatic character.

Little of the material is fully novel. The work echoes themes raised several years ago by [Ian Kerr and Carissima Mathen](#) and extended more recently by [Rebecca Crootof](#), among others. Its elegance lies in the coordination of prior thinking and writing in an unusually clear way. The framework can be applied generally to the roles that algorithms increasingly play in governance of many sorts, from urban planning to professional sports.

I’ll close with an illustration of that point, one that appears mostly, and briefly, in the footnotes. Consider soccer, or football, as it is known in much of the world. Re and Solow-Niederman acknowledge the utility of thinking about sports as a case study with respect to automation and adjudication. (P. 254 n. 37; P. 278 n. 121.) The following picks up on this and extends it, to show how their framework can be applied to help clarify thinking about a specific example. Other scholars have done similar work, notably [Meg Jones and Karen Levy in *Sporting Chances: Robot Referees and the Automation of Enforcement*](#). But they did not include soccer in their group of cases, and automation in soccer refereeing has some distinctive attributes that may be particularly relevant here.

A few years ago, to improve refereeing in professional football matches, VAR systems (short for Video Assistant Referee) were introduced. During breaks in play, referees are permitted to look at recorded video of the game and consult with off-field officials who supervise video playback.

VAR has been controversial. It has been implemented so far in a “division of labor” sense, against a long history of experimentation with the rules of the game (or [“laws,” as they are formally known](#)). VAR data are generally determinative with respect to rule-based judgments, such as whether a goal has been scored. VAR data are employed differently with respect to possible penalty kicks and possible ejections. In both contexts, presumably because of the severity of the consequences (or, perhaps, despite them), VAR data are advisory. The human referee retains the discretion to make final determinations.

The relevance of VAR is not its technical details; the point is its systems impact. As Jones and Levy note, a mechanical element has been introduced in a game in which both play and adjudication have long been inescapably error-prone. Rightness and wrongness, even in a yes/no sense, are human and humane constructs, in soccer and in the law. VAR, like an AI judge, changes something about this human “essence” of playing and judging experiences.

But the VAR example illuminates something critical about Re and Solow-Neiderman’s framework. Soccer referees not only adjudicate yes/no applications of the rules. Penalty kicks and player ejections do not follow only from administration of soccer’s laws in a “correct/incorrect” sense, with accuracy as the paramount value. In the long history and narrative of soccer, the referee’s discretion has always represented justice. Does a violent tackle warrant a penalty kick? Sometimes it does; sometimes it does not. Unlike referees’ decisions in other sports with machine-based officiating, critical judgments in soccer are based on “fairness” rather than only on “the rule,” where “fairness” is equated to a sense of earned outcomes, or “just deserts.” The soccer referee is dispenser of what might be called “equitable justice” on the field. Enlisting VAR risks tilting this decision-making process toward what might be called “codified justice.”

Is this good for the game, or for the society that depends on it? It’s too soon to say. Soccer, like all institutions, has never been unchanging. Soccer laws, soccer technologies, and soccer values are always at least a little bit in flux, and sometimes much more so. But the soccer example offers not simply another way of understanding challenges of AI and the law. Re and Solow-Neiderman have given us a framework based in the law that helps us understand the challenges of automation and algorithms across additional critical domains of social life. Those challenges ask us to consider, again, what we mean by justice—not only in the law but also beyond it.

Cite as: Michael Madison, *Oyez! Robot*, JOTWELL (January 24, 2020) (reviewing Richard M. Re & Alicia Solow-Niderman, *Developing Artificially Intelligent Justice*, 22 **Stan. Tech. L. Rev.** 242 (2019)), <https://cyber.jotwell.com/oyez-robot/>.