

Automatic – for the People?

Author : Tal Zarsky

Date : November 8, 2016

Andrea Roth, [Trial by Machine](#), 104 **Georgetown Law Journal** 1245 (2016).

Crucial decision-making functions are constantly migrating from humans to machines. The criminal justice system is no exception. In a recent insightful, eloquent, and rich [article](#), [Professor Andrea Roth](#) addresses the growing use of machines and automated processes in this specific context, critiquing the ways these processes are currently implemented. The article concludes by stating that humans and machines must work in concert to achieve ideal outcomes.

Roth's discussion is premised on a rich historical timeline. The article brings together measures old and new—moving from the polygraph to camera footage, impairment-detection mechanisms such as Breathalyzers, and DNA typing, and concluding with AI recommendation systems of the present and future. The article provides an overall theoretical and doctrinal discussion and demonstrates how these issues evolved. Yet it also shows that as time moves forward, problems often remain the same.

The article's main analytical contribution is its two central factual assertions: First, that machines and mechanisms are introduced unequally, as a way to strengthen the prosecution and not to exonerate. In other words, there are no similar opportunities to apply these tools to enhance defendants' cases. Secondly, machines and automated processes are inherently flawed. This double analytic move might bring a famous ["Annie Hall" joke](#) to mind: "The food at this place is really terrible . . . and such small portions."

The article's first innovative and important claim—regarding the pro-prosecution bias of decisions made via machine—is convincing and powerful. Roth carefully works through technological test cases to show how the state uses automated and mechanical measures to limit "false negatives"—instances in which criminals eventually walk free. Yet when the defense suggests using the same measures to limit "false positives"—the risk that the innocent are convicted—the state pushes back and argues that machines and automated processes are problematic. Legislators and courts would be wise to act upon this critique and consider balancing the usage of automated measures.

Roth's second argument—automation's inherent flaws—constitutes an important contribution to a growing literature pointing out the problems of automated processes. The article explains that such processes are often riddled with random errors which are difficult to locate. Furthermore, they are susceptible to manipulation by the machine operators. Roth demonstrates in several contexts how subjective assumptions can be and are buried in code, inaccessible to relevant litigants. Thus, the so-called "objective" automated process in fact introduces unchecked subjected biases of the system's programmers. Roth further notes that the influence of these biased processes is substantial. Even in instances in which the automated processes are intended to merely recommend an outcome, the humans using it give extensive deference to the automated decision.

The article fairly addresses counter-arguments, noting the virtues of automated processes. Roth explains how automated processes can overcome systematic human error and thus limit false positives in the context of DNA evidence and computer-assisted sentencing. To this I might add that machines allow for replacing decisions made in the periphery of systems with those made by central planners. In many instances, it might be both efficient and fair to prefer systematic errors made by the central authority to the biases arising when rules are applied with discretion in the field and subjected to the many biases of agents.

In addition, Roth explains that automated processes are problematic, as they compromise *dignity*, *equity*, and *mercy*. Roth's argument that trial by machine compromises dignity is premised on the fact that applying some of these mechanical and automated measures calls for degrading processes and the invasion of the individual's property.

This dignity-based argument could have been strengthened by a claim often voiced in Europe: to preserve dignity, a human should be subjected to the decision of a fellow human, especially when there is much at stake. Anything short of that will prove to be an insult to the affected individual's honor. Europeans provide strong legal protections for dignity which are important to mention—especially given the growing influence of EU law (a dynamic at times referred to as the "[Brussels Effect](#)"). Article 22 of the recently introduced General Data Protection Regulation ([GDPR](#)) provides that individuals have the right not to be subjected to decisions that are "based solely on automated processing" when these are deemed to have a significant effect. Article 22 provides several exceptions, yet individuals must be provided with a right to "obtain human intervention," and have the ability to contest the automated findings and conduct additional examinations as to how the decision was reached (see also Recital 71 of the GDPR). Similar provisions were featured in Article 12(a) and Article 15 of the [Data Protection Directive](#) which the GDPR is set to replace over the next two years, and in older French legislation. To be fair, it is important to note that in some EU Member States these provisions have become dead letters. Their recent inclusion in the GDPR will no doubt revive them. However, the GDPR does not pertain to criminal adjudication.

Roth's argument regarding equity (or the lack thereof in automated decisions) is premised on the notion that automated processes are unable to exercise moral judgment. Perhaps this is about to change. Scholars are already suggesting the [creation of automated tools that will do precisely that](#). Thus, this might not be a critique of the processes in general, but of the way they are currently implemented—a concern that could be mitigated over time as technology progresses.

The lack of mercy in machine-driven decisions is obviously true. However, the importance of building mercy into our legal systems is debatable. Is the existing system equally merciful to all social segments? One might carefully argue that very often the gift of mercy is yet another privilege of the elites. As I [argue elsewhere](#), automation can remove various benefits the controlling minorities still have—such as the cry for mercy—and this might indeed explain why societies are slow to adopt these measures, given the political power of those to be harmed from its expansion.

To conclude, let's return to Woody Allen and the "Annie Hall" reference. If, according to Roth, automated processes are problematic, why nonetheless should we complain that the portions are so small, and consider expanding their use to limit "false positives"? Does making both claims make sense? I believe it does. For me and others who are unconvinced that automated processes are indeed problematic (especially given the alternatives) the article both describes a set of problems with automation we must consider, and also provides an alarming demonstration of the injustices unfolding in implementation. But joining these two arguments should also make sense to those already convinced that machine-driven decisions are highly problematic. This is because it is quite clear that machines and automated processes are here to stay. Therefore, it is important both to identify their weaknesses and improve them (at times by integrating human discretion) and to assure that the advantages they provide are equally shared throughout society.

Cite as: Tal Zarsky, *Automatic – for the People?*, JOTWELL (November 8, 2016) (reviewing Andrea Roth, *Trial by Machine*, 104 *Georgetown Law Journal* 1245 (2016)), <http://cyber.jotwell.com/automatic-for-the-people/>.