Military Algorithms and the Virtues of Transparency

Author : Kristen Eichensehr

Date : November 20, 2019

Ashley S. Deeks, Predicting Enemies, 104 **Va. L. Rev.** 1529 (2018).

For all the justifiable concern in recent years directed toward the prospect of autonomous weapons, other military uses of automation may be more imminent and more widespread. In *Predicting Enemies*, Ashley Deeks highlights how the U.S. military may deploy algorithms in armed conflicts to determine who should be detained and for how long, and who may be targeted. Part of the reason Deeks predicts these near-term uses of algorithms is that the military has models: algorithms and machine-learning applications currently used in the domestic criminal justice and policing contexts. The idea of such algorithms being employed as blueprints may cause heartburn. Their use domestically has prompted multiple lines of critique about, for example, biases in data and lack of transparency. Deeks recognizes those concerns and even intensifies them. She argues that concerns about the use of algorithms are exacerbated in the military context because of the "double black box"—"an 'algorithmic black box' inside what many in the public conceive of as the 'operational black box' of the military" (P. 1537)—that hampers oversight.

Predicting Enemies makes an important contribution by combining the identification of likely military uses of algorithms with trenchant critiques drawn from the same sphere as the algorithmic models themselves. Deeks is persuasive in her arguments about the problems associated with military deployment of algorithms, but she doesn't rest there. She argues that the U.S. military should learn from the blowback it suffered after trying to maintain secrecy over post-9/11 operations, and instead pursue "strategic transparency" about its use of algorithms. (P. 1587.) Strategic transparency, as she envisions it, is an important and achievable step, though likely still insufficient to remedy all of the concerns with military deployment of algorithms.

Deeks highlights several kinds of algorithms used domestically and explains how they might parallel military applications. Domestic decision-makers use algorithms to assess risks individuals pose in order to determine, for example, whether to grant bail, impose a prison sentence, or allow release on parole. Even more controversially, police departments use algorithms to "identif[y] people who are most likely to be party to a violent incident" in the future (P. 1543, emphasis omitted), as well as to pinpoint geographic locations where crimes are likely to occur.

These functions have military counterparts. During armed conflicts, militaries often detain individuals and have to make periodic assessments about whether to continue to detain them based on whether they continue to pose a threat or are likely to return to the fight. Militaries, like police departments, also seek to allocate their resources efficiently. Algorithms that predict where enemy forces will attack or who is likely to do the attacking, especially in armed conflicts with non-state armed groups, would have obvious utility.

But, Deeks argues, problems with domestic use of algorithms are exacerbated in the military context. As compared with domestic police departments or judicial officials, militaries using algorithms early in a particular conflict are likely to have far less and less granular information about the population with which to train their algorithms. And algorithms trained for one conflict may not be transferable to different conflicts in different locations involving different populations, meaning that the same problems with lack of data would recur at the start of each new conflict. There's also the problem of applying algorithms "cross-culturally" in the military context, rather than "within a single society" as is the case when they are used domestically (P. 1565), and the related possibility of exacerbating biases embedded in the data. With bad or insufficient data come inaccurate algorithmic outcomes.

Deeks also worries about "automation bias"—that military officials will be overly willing to trust algorithmic outcomes

and even more susceptible to this risk than judges, who are generally less tech-savvy. (Pp. 1574-75.) At the same time, she also warns that a lack of transparency about how algorithms work could make military officials unwilling to trust algorithms when they should, that is, when the algorithms would actually improve decision-making and compliance with international law principles like distinction and proportionality. (Pp. 1568-71.)

These and other concerns lead Deeks to her prescription for "strategic transparency." Deeks argues that the military should "fight its institutional instincts" (P. 1576) to hide behind classification and limited oversight from Congress and the public and instead deploy a lesson from the war on terror—that "there are advantages to be gained by publicly confronting the fact that new tools pose difficult challenges and tradeoffs, by giving reasons for their use, and by clarifying how the tools are used, by whom, and pursuant to what legal rules." (P. 1583.) Specifically, Deeks argues that in pursuing transparency, the military should explain when and how it uses algorithms and machine learning, articulate how such tools comply with its international law obligations, and engage in a public discussion of costs and benefits of using algorithms. (Pp. 1588-89.) She also urges the military to "articulate[] how it will test the quality of its data, avoid training its algorithms on biased data, and train military users to avoid falling prey to undue automation biases." (P. 1590.)

Deeks previously served as the Assistant Legal Adviser for Political-Military Affairs in the State Department's Office of the Legal Adviser (where I had the pleasure of working with her), and so she has the experience of an internal advisor combined with the critical eye of an academic commentator. One hopes that the U.S. military—and others around the world—will heed her thoughtful advice about transparency in the use of algorithms. Transparency is not a panacea for problems of data availability, quality, and bias, but it may help with oversight and accountability. And that's a good first step.

Cite as: Kristen Eichensehr, *Military Algorithms and the Virtues of Transparency*, JOTWELL (November 20, 2019) (reviewing Ashley S. Deeks, *Predicting Enemies*, 104 **Va. L. Rev.** 1529 (2018)), https://cyber.jotwell.com/military-algorithms-and-the-virtues-of-transparency/.