Automation, Trust, Responsibility in Algorithmic Warfare

Stefka Hristova, Michigan Technological University (United States)

International Conference on Computer Ethics: Philosophical Enquiry (CEPE) 2023, Chicago, IL

Keywords: war, algorithms,action, trust, responsibility, autonomy, plausible deniability, unmannedness

Extended Abstract

In a 2006 editorial for National Defense, Stew Magnuson made an apt observation: "The robot army is coming." Algorithmic war has been envisioned as war fought by algorithmic technology under the guise of protecting human life and in response to a potential enemy robot army. As David Humbling has reported, in preparation for this new war, "[o]ne U.S. Navy project envisages having to counter up to a million drones at once" (Hambling 2021). The algorithmic technology developed is indeed one that envisions both attacks and counterattacks as air combat. The military's robot army increasingly consists of autonomous technology deployed on jets and drones. In 2020, the "U.S. Air Force let an artificial intelligence take over the navigation and sensor systems of a Lockheed U-2 spy jet during a training flight [marking] the first known time an AI has to been used to control a US military aircraft" (The Airforce 2020). Here, onboard the U-2 "Dragon Lady" spy plane, the "human Air Force officer" was partnered with "ARTUμ algorithm" which is now responsible for real-world sensor monitoring and navigation and yet is modeled after a gaming system (Browne 2020). While these seem like small, incremental steps toward algorithmic war, they point to an ambitious goal where in "10 to 15 years max, you are going to see the widespread, ubiquitous use of robots throughout most militaries in the world" (2020). This idea of robot-driven warfare has been met with skepticism as it raises significant moral and ethical issues about trust and responsibility.

Trust

War systems are increasingly seen as entirely unmanned and thus autonomous. The processes of automation of war require the articulation of three major interrelated processes as they relate to trust. As Paul Scharre has aptly written, "Activating an autonomous system is an act of trust" (2018, 149). First, the process of building trust in human-machine partnerships and then building trust in the machine algorithms themselves. Second, trust needs to be established in relation to the amount of error or risk that an algorithm is allowed to accept. Autonomous technology is also a system of risk. "The key factor to assess with autonomous systems isn't whether the system is better than a human, but rather if the system fails (which it inevitably will), what is the amount of damage it could cause, and can we live with that risk" (193)? Third, trust figures into relegating the ethical and moral responsibility of warfare away from human agents and onto autonomous technologies. It is important to note that these processes are biopolitical and that the conversation about automation only addresses the side firing the guns. The victims of warfare remain vulnerable and also human.