

Ethics Memo #2

“Guns don't do the killing, people do.” Well, what if artificial intelligence is holding the gun? Are people still responsible for the killing? Immense advances in remotely-piloted aircraft and artificial intelligence (AI) have been made by engineers contracted with the Defense Advanced Research Projects Agency (DARPA). In August of 2020, a contracted AI agent successfully defeated an F-16 fighter pilot in a series of simulated combat engagements during the final round of DARPA's “Alpha Dogfight” Trials.

This technology is capable of selecting and engaging targets without human control. A fighter plane equipped with artificial intelligence could eventually execute tighter turns, take greater risks, and even get off better shots than human pilots. The Department of Defense continues to conduct research and implement projects that will one day incorporate artificial intelligence into warfighting.

Implementing AI into fighter jets poses many ethical concerns. The rest of this memo will focus on three key ethical questions: humanness, error, and judgment.

War is already an ethically contested subject. Administering this type of technology to fighter jets would only aid in the inhumanity of warfare. By replacing a soldier's job with a computer, we completely detach human emotion and empathy away from warfare. Removing the human experience from combat would most likely result in more destruction and vicious acts. Think of social media for example. According to KQED, “being online lowers your inhibitions. This often

results in people either behaving meaner or opening up more online than they normally would in face-to-face conversations.” In class, we discussed how our interactions over the internet are “disembodied.” Those who are terminally online lose all feelings of shame. If it is human nature to act more “bold” or “mean” over the internet, who is to say we would not have the same reaction to AI use in combat. It is a plausible fear that soldiers will cause more destruction using AI because they no longer have to confront the situation face-to-face. In other words, combat would become a disembodied experience. This causes me to question whether we would lose all humanness in warfare by implementing AI into fighter jets.

There are no technological designs that are perfect. In fact, according to IEEE “the neutral network technology that drives many AI systems can break down in ways that remain a mystery to researchers.” If AI is intrinsically untrustworthy, we should never feel comfortable implementing it into our military. What would happen if AI were to kill our own soldiers? Murder innocent bystanders? Or even destroy historical artifacts? Who would be at fault? Who would impose reparations? There is simply too high of a risk when AI should fail in combat. In addition, it is proven that AI can be inaccurate, culturally insensitive, and neurodivergent. When human lives are at stake, we should not put our trust in artificial intelligence.

According to the Oxford Dictionary, artificial intelligence is “...able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.” In warfare, split-second decision-making and ethical judgment are paramount. By implementing AI into fighter jets we will not only dehumanize warfare but also completely eliminate judgment. There is no formula for war. In a

situation where plans change, a human soldier would adapt, and an AI fighter jet would continue to shoot. AI replacing human soldiers is dangerous because AI does not have the ability to pass judgment on a situation, it will simply do what it has been programmed to do.

There is no technological design that is perfect. All technology has the potential to cause suffering. To determine if a certain technology is safe for public use, it is the duty of engineers to determine whether its potential for good outweighs the bad. In the case of AI use in fighter jets, the bad outweighs the good. My solution to this problem is to use AI as a way to train our soldiers, rather than using AI to replace them. This way, when AI inevitably fails, the consequences are no longer catastrophic. In addition, humanness and judgment will be restored with the implementation of human pilots.

A reasonable objection is that AI will not replace pilots, rather it will work as a “co-pilot.” As DARPA envisions it, “AI will fly the plane in partnership with the pilot, who will remain ‘in the loop,’ monitoring what the AI is doing and intervening when necessary” (New Yorker).

Unfortunately, this objection is still unethical. Even if there is a human present to pass judgment, it is still a computer that is doing the killing. Furthermore, how will soldiers determine when to interfere with the AI? When AI inevitably fails, will the soldier “co-piloting” be at fault? There are too many ethical questions that arise when it comes to AI use in combat. Even if AI were to work as a “co-pilot” the bad still outweighs the good. The only plausible solution is to use this technology solely for training purposes.

SOURCES:

Choi, C. Q. (2021, September 21). 7 Revealing Ways AIs Fail. IEEE Spectrum; IEEE Spectrum. <https://spectrum.ieee.org/ai-failures>

Is the Internet Making You Meaner? | KQED Education. (2019, August 5). KQED.
<https://www.kqed.org/education/532334/is-the-internet-making-you-meaner>

Halpern, S. (2022, January 14). The Rise of A.I. Fighter Pilots. The New Yorker; The New Yorker.
<https://www.newyorker.com/magazine/2022/01/24/the-rise-of-ai-fighter-pilots>

Heins, John C, & Naval. (2018). Airpower: The Ethical Consequences of Autonomous Military Aviation. DTIC. <https://apps.dtic.mil/sti/citations/AD1079772>

What, A. (2017, March). *Artificial Intelligence (AI): What is it and how does it work?* Lexology; ClaimVantage.
<https://www.lexology.com/library/detail.aspx?g=5424a424-c590-45f0-9e2a-ab05daf032d#:~:text=The%20Oxford%20Dictionary%20defines%20Artificial,making%2C%20and%20translation%20between%20languages.>