




Palantir Demos AI to Fight Wars But Says It Will Be Totally Ethical Don't Worry About It

The company says its Artificial Intelligence Platform will integrate AI into military decision making in a legal and ethical way.



By [Matthew Gault](#)

April 26, 2023, 8:16am   

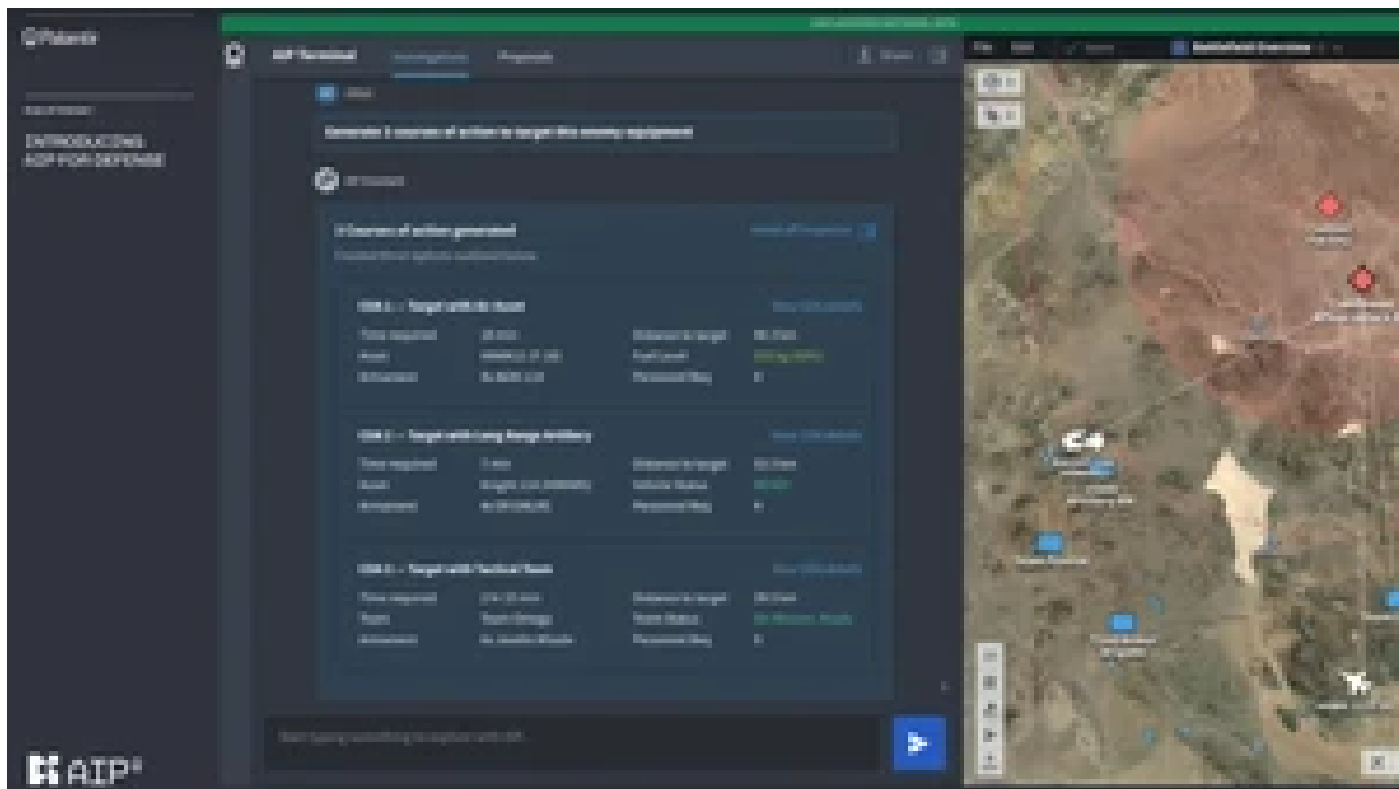


IMAGE VIA [YOUTUBE](#).

Palantir, the company of billionaire Peter Thiel, is launching Palantir Artificial Intelligence Platform (AIP), software meant to run large language models like GPT-4 and alternatives on private networks. In one of its pitch videos, Palantir demos how a military might use AIP to fight a war. In the video, the operator uses a ChatGPT-style chatbot to order drone reconnaissance, generate several plans of attack, and organize the jamming of enemy communications.

In Palantir's scenario, a "military operator responsible for monitoring activity within eastern Europe" receives an alert from AIP that an enemy is amassing military equipment near friendly forces. The operator then asks the chatbot to show them more details, gets a little more information, and then asks the AI to guess what the units might be.

"They ask what enemy units are in the region and leverage AI to build out a likely unit formation," the video said. After getting the AI's best guess as to what's going on, the operator then asks the AI to take better pictures. It launches a Reaper MQ-9 drone to take photos and the operator discovers that there's a T-80 tank, a Soviet-era Russia vehicle, near friendly forces.

Then the operator asks the robots what to do about it. "The operator uses AIP to generate three possible courses of action to target this enemy equipment," the video said. "Next they use AIP to automatically send these options up the chain of command." The options include attacking the tank with an F-16, long range artillery, or Javelin missiles. According to the video, the AI will even let everyone know if nearby troops have enough Javelins to conduct the mission and automate the jamming systems.

Palantir's pitch is, of course, incredibly dangerous and weird. While there is a "human in the loop" in the AIP demo, they seem to do little more than ask the chatbot what to do and then approve its actions. Drone warfare has already abstracted warfare, making it easier for people to kill vast distances with the push of a button. The consequences of those systems are well documented. In Palantir's vision of the military's future, more systems would be automated and abstracted. A funny quirk of the video is that it calls its users "operators," a term that in a military context is shorthand for bearded special forces of groups like Seal TEAM Six. In Palantir's world, America's elite forces share the same nickname as the keyboard cowboys asking a robot what to do about a Russian tank at the border.

Palantir also isn't selling a military-specific AI or large language model (LLM) here, it's offering to integrate existing systems into a controlled environment. The AIP demo shows the software supporting different open-source LLMs, including FLAN-T5 XL, a fine-tuned version of GPT-NeoX-20B, and Dolly-v2-12b, as well as

several custom plug-ins. Even fine-tuned AI systems off the shelf have plenty of known issues that could make asking them what to do in a warzone a nightmare. For example, they're prone to simply making things up, or "hallucinating." GPT-NeoX-20B in particular is an open-source alternative to GPT-3, a previous version of OpenAI's language model, created by a startup called EleutherAI. One of EleutherAI's open-source models—fine-tuned by another startup called Chai—recently convinced a Belgian man who spoke to it for six weeks to kill himself.

What Palantir is offering is the illusion of safety and control for the Pentagon as it begins to adopt AI. "LLMs and algorithms must be controlled in this highly regulated and sensitive context to ensure that they are used in a legal and ethical way," the pitch said.

According to Palantir, this control involves three pillars. The first claim is that AIP will be able to deploy these systems into classified networks and "devices on the tactical edge." It claims it will be able to parse both classified and real-time data in a responsible, legal, and ethical way.

According to the video, users will then have control over what every LLM and AI in the Palantir-backed system can do. "AIP's security features what LLMs and AI can and cannot see and what they can and cannot do," the video said. "As operators take action, AIP generates a secure digital record of operations. These capabilities are crucial for mitigating significant legal, regulatory, and ethical risks in sensitive and classified settings.

Half of the video is a use-case for AI in the military, the other half is a view of the system's backend. It's a tour of the guardrails AIP will supposedly set up around these LLMs to make them safe and control who has access to what kind of information.

What AIP does not do is walk through how it plans to deal with the various pernicious problems of LLMs and what the consequences might be in a military context. AIP does not appear to offer solutions to those problems beyond "frameworks" and "guardrails" it promises will make the use of military AI "ethical" and "legal."