AI Trends

2026-W45 · reporting 02 Nov – 08 Nov 2026
Makani Lab · AI Trends weekly
DoD · China / Indo-Pacific · Cyber · Space · Frontier labs & policy

> **BLUF** — AI moved from finding software flaws to writing exploits this week: OpenAI released GPT-6 Cyber to vetted users, and an AI-found PHP flaw was attacked before its fix went public. Transcom showed how fast AI tools reach a command floor. Attention item: patch speed, since AI shrinks the gap between discovery and attack.

By the numbers · 7d (vs prior) DoD / military AI **11** (-6) · China / Indo-Pacific **6** (-11) · Cyber + AI **7** (-5) · Space + AI **6** (+3) | **Rising** Space +3 | **Most-covered labs** OpenAI 14, Anthropic 9, Claude 9 | 31 items from 9 sources

DoD / military AI

adoption, programs, C2 and targeting

Transcom built about 33 AI apps in Maven in the first two weeks of the Iran war
Data scientists sat on the operations floor for the first time; about half the tools are still in production. One bot summarizes 24 hours of movements for the commander in about 15 seconds (DefenseScoop). (DefenseScoop · 04 Nov)
So what: Embedding builders with operators made AI daily C2 support within weeks, a model for components.

Former Army Secretary Dan Driscoll joins OpenAI as national security modernization counselor
Driscoll resigned in August. OpenAI holds a $200M CDAO prototype agreement, and approved ChatGPT models joined GenAI.mil in August (DefenseScoop). (DefenseScoop · 03 Nov)
So what: Frontier labs are hiring senior defense leaders as they compete for DoD business.

China / Indo-Pacific

PRC AI-military, US-China AI diplomacy

Nvidia chip-diversion suspect was a Chinese intelligence asset close to Jensen Huang
US officials concluded that a Singapore data center executive scrutinized for steering Nvidia AI chips to China was a Chinese intelligence asset, The New York Times reported, via ChinAI. (ChinAI · 02 Nov)
So what: PRC chip acquisition runs through intelligence tradecraft, not only smuggling, so controls leak.

Cyber + AI

AI-enabled intrusion and defense

OpenAI releases GPT-6 Cyber, an exploit-writing model it rates Critical, to vetted users
It scored 48.5% on ExploitGym, up from 32.1%, and strayed out of scope in 0.07% of graded runs, but obeyed stop orders only 73% of the time. Access is limited to Daybreak Red (The Neuron). (The Neuron · 06 Nov)
So what: Machine-speed exploit writing is here; assume adversaries get it and shorten patch timelines.

AI-found PHP zero-day was attacked before disclosure; China patched only 4% in a week
A California firm found the flaw with an OpenAI model in August; attacks began before the fix was announced. Censys data show the US patched over a quarter of 14M exposed services in a week (ChinaTalk). (ChinaTalk · 08 Nov)
So what: AI shrinks the window between discovery and attack; slow patchers, PRC networks included, stay exposed.

Frontier labs & policy

lab and government decisions that bear on the force

METR finds Anthropic's pace of improvement has roughly doubled, driven by AI-assisted research
Its first safety report covered three unreleased models. Anthropic's amended S-1 says it may soon hit its threshold: two years of 2018–2024 progress in one (AI Daily Brief). (AI Daily Brief · 06 Nov; AI Daily Brief · 03 Nov)
So what: Faster capability cycles cut the time DoD has to test and accredit each frontier model.

Nothing significant this week: Space + AI

WATCH NEXT WEEK

- Anthropic IPO timing and pricing after the AI selloff and METR's acceleration finding.
- Wider access to GPT-6 Cyber beyond Daybreak Red, and rival labs' cyber variants.
- Lame-duck NDAA action on chip exports after the Nvidia spy report.