The AI Cold War's Newest Weapon: Stealing the Teacher's Answer Key

Date: May 03, 2026 | Model: anthropic-batch:claude-opus-4-7

Source: PDF: White_House_accuses_China_of_'industrial_scale'_theft_of_AI_technology.pdf

Contents

Note: the original article is provided as a separate file (attached to the email or downloadable from the website).

1. Explanation

> ***Imagine cheating on a test by secretly photographing every answer the smartest kid writes — then selling your knockoff version for a tenth of the price. That's what the White House just accused China of doing to American AI.***

What's Going On?

On April 23, 2026, the White House accused Chinese entities of running 'industrial-scale' campaigns to steal from American AI labs through a technique called distillation. In a memo, Michael Kratsios — the chief science and technology adviser to the president — said foreign actors are using thousands of fake accounts and 'jailbreaking' tricks to extract knowledge from US frontier AI models like those built by OpenAI and Anthropic. Anthropic, in a separate February 2026 disclosure, said it had documented more than 24,000 such fraudulent accounts attached to three Chinese AI firms.

The accusation lands weeks before Donald Trump is scheduled to meet Xi Jinping in Beijing. China's embassy called the claims 'pure slander.' Meanwhile, Congress passed bills that could blacklist Chinese AI firms suspected of distillation, cutting them off from US technology entirely.

How To Think About It

Distillation is a normal AI technique — it's how companies make smaller, cheaper models by training them to mimic bigger ones. The fight is over who's allowed to do it to whom.

- Think of it like sneaker resale: Nike designs the Jordan, but if a factory in another country reverse-engineers it and sells near-identical pairs for $30, Nike loses the value of its R&D investment.
- Or think of the Cold War's nuclear race: the US spent billions developing the atomic bomb, then Soviet spies copied the blueprints in years. The lead a country thinks it has can evaporate fast when secrets leak.

Key Things To Know

- DeepSeek, a Chinese AI firm, shocked the industry in early 2025 by releasing a model nearly as capable as OpenAI's at a fraction of the cost — OpenAI claims it used GPT outputs in training, violating its terms of service.
- Distillation works by feeding millions of questions to a powerful model and using its answers to train a smaller copycat — like learning chess by watching a grandmaster's every move.
- The US has restricted China's access to advanced Nvidia AI chips since 2022, hoping to slow it down; distillation is one workaround that requires far less computing power.
- Anthropic has accused three Chinese firms — DeepSeek, Moonshot and MiniMax — of running over 16 million distillation exchanges via 24,000+ fake accounts; the 'entity list' blacklist could now lock them out of US tech entirely.
- What most people miss: the copycat models often lack safety guardrails, meaning a stripped-down clone could more easily help build bioweapons or launch cyberattacks.

Why It Matters

If you're considering a career in tech, computer science or international policy, this is the defining conflict of the next two decades. AI capability is now treated like nuclear capability was in 1955 — a strategic asset governments will fight, sanction and spy over. The apps you use, the chips in your phone, even which universities can collaborate with whom: all of it is being reshaped by this rivalry right now.

The Bigger Picture

The deeper question is whether AI leadership can be defended at all. Knowledge tends to leak — through people, papers, products and now apparently through models talking to each other. If distillation is unstoppable, the US edge built on expensive chips and elite researchers may be smaller than Washington thinks. Watch for: tighter export controls, possible bans on Chinese users accessing US AI services, and a fragmenting global internet where American and Chinese AI ecosystems no longer talk to each other at all.

2. Key Terms Glossary

Distillation

An AI training method where a smaller model learns by mimicking the outputs of a larger, more powerful one — useful when done legitimately, controversial when done by copying someone else's proprietary model.

Frontier AI

The most advanced AI systems currently in existence, like GPT-class models, that push the limits of what the technology can do.

Jailbreaking

Tricking an AI model into bypassing its built-in restrictions — for example, getting it to reveal information or behaviour its creators tried to block.

Entity list

A US government blacklist of foreign companies and organisations that American firms are largely banned from selling technology to; being added to it can cripple a company's supply chain.

Export controls

Government rules restricting which technologies can be sold abroad — the US uses them to prevent China from acquiring advanced AI chips and equipment.

Proxy accounts

Fake or borrowed user accounts used to disguise who is really accessing a service, often to evade detection or rate limits.

Intellectual property

Legally protected creations of the mind — patents, trade secrets, copyrighted code — that companies and countries treat as valuable assets worth defending.

Safeguards (in AI)

Built-in restrictions that prevent a model from producing dangerous content, such as instructions for weapons or malware.

3. Reading Comprehension Quiz

Circle the best answer for each question.

Q1. Which choice best states the central idea of the passage?

- ◯ **A)** China has officially admitted to copying American AI models for military use.
- ◯ **B)** The US accuses China of large-scale AI theft and is preparing countermeasures.
- ◯ **C)** Distillation is an illegal technique that no AI company should ever be using.
- ◯ **D)** President Trump plans to cancel his upcoming meeting with President Xi Jinping.
- ◯ **E)**

Q2. According to the passage, why are American AI companies concerned about distilled models specifically on national security grounds?

- ◯ **A)** Because distilled models are more powerful than the originals they copy.
- ◯ **B)** Because distilled models are sold at higher prices to foreign militaries.
- ◯ **C)** Because distilled models often lack safeguards against dangerous content.
- ◯ **D)** Because distilled models cannot be detected by US intelligence agencies.
- ◯ **E)**

Q3. As used in the passage, the word 'distil' most nearly means

- ◯ **A)** to purify a liquid by heating it
- ◯ **B)** to extract knowledge by mimicking outputs
- ◯ **C)** to summarise a long document briefly
- ◯ **D)** to destroy a competitor's product
- ◯ **E)**

Q4. As used in the passage, 'frontier' most nearly means

- ◯ **A)** the border between two countries
- ◯ **B)** an unsettled wilderness region
- ◯ **C)** the most advanced or cutting-edge
- ◯ **D)** a dangerous and unstable territory
- ◯ **E)**

Q5. Which statement about export controls can most reasonably be inferred from the passage?

- ◯ **A)** Export controls have completely stopped China's progress in AI development.
- ◯ **B)** Export controls were created specifically to address distillation attacks.
- ◯ **C)** Distillation may partially undermine the advantage that export controls create.
- ◯ **D)** Export controls are universally supported by American AI company executives.
- ◯ **E)**

Q6. The passage suggests that DeepSeek's emergence was significant primarily because

- ◯ **A)** it proved Chinese firms could match US capabilities much more cheaply
- ◯ **B)** it was the first AI company to operate without any government support
- ◯ **C)** it openly admitted to violating OpenAI's terms of service publicly
- ◯ **D)** it developed entirely new AI techniques unknown to American labs
- ◯ **E)**

Q7. The author's tone in describing the White House accusations is best described as

- ◯ **A)** openly skeptical of the US government's claims
- ◯ **B)** neutrally reporting both accusations and Chinese denials
- ◯ **C)** enthusiastically supportive of harsh sanctions against China
- ◯ **D)** deeply alarmed about an imminent military confrontation
- ◯ **E)**

Q8. The author's primary purpose in this article is to

- ◯ **A)** argue that the United States must immediately ban Chinese AI firms
- ◯ **B)** explain a new escalation in US-China tensions over AI technology
- ◯ **C)** warn readers that AI bioweapons are about to be deployed worldwide
- ◯ **D)** celebrate American technological superiority over Chinese competitors
- ◯ **E)**

Q9. Which statement about the relationship between computing power and distillation can most reasonably be inferred?

- ◯ **A)** Distillation requires significantly more computing power than original training.
- ◯ **B)** Computing power is irrelevant to whether AI models can be distilled.
- ◯ **C)** Distillation offers a way around limited access to powerful computing chips.
- ◯ **D)** Chinese firms now have more computing power than American AI labs do.
- ◯ **E)**

Q10. Which choice provides the best evidence for the answer to the previous question?

- ◯ **A)** 'The US and China are engaged in an arms race for AI technology.'
- ◯ **B)** 'Chinese AI firms are relying on distillation attacks to offset deficits in AI computing power.'
- ◯ **C)** 'Models created by surreptitious, unauthorised distillation campaigns did not match the performance of the original models.'
- ◯ **D)** 'The House Foreign Affairs Committee on Wednesday passed a slew of bills.'
- ◯ **E)**

My Score: _____ / 10

4. Answer Key with Explanations

Q1. Which choice best states the central idea of the passage?

Answer: B

B captures the article's core: the White House memo, the accusations, and the policy response. A is false — China denies the claims. SAT Tip: For central-idea questions, the right answer usually summarises both the situation AND the response or stakes, not just one half of the story.

Q2. According to the passage, why are American AI companies concerned about distilled models specifically on national security grounds?

Answer: C

The passage explicitly states that distilled models 'lack the safeguards that, for example, prevent the development of bioweapons.' A is wrong (Trap A — opposite direction; the article says distilled models do NOT match original performance). SAT Tip: When a question asks 'why,' scan the passage for cause-effect language like 'because,' 'enables' or 'pose' — the answer is usually within one or two sentences of that signal word.

Q3. As used in the passage, the word 'distil' most nearly means

Answer: B

In context, 'distil' refers to training smaller models from larger ones' outputs — extracting capability. A is the everyday chemistry meaning (Trap B — common usage, wrong context). SAT Tip: On vocab-in-context, substitute each option into the original sentence — the right answer keeps the sentence's overall meaning intact.

Q4. As used in the passage, 'frontier' most nearly means

Answer: C

'Frontier AI systems' refers to the most advanced models currently being developed. A and B are common geographic meanings (Trap B — passage uses a metaphorical sense). SAT Tip: Technical fields often borrow ordinary words and give them precise new meanings — trust the surrounding sentence over your prior associations.

Q5. Which statement about export controls can most reasonably be inferred from the passage?

Answer: C

The passage notes US firms worry distillation lets foreign labs 'close the competitive advantage' that export controls create — implying controls work partly, but distillation chips away at them. A is too absolute (Trap A — wrong scope). SAT Tip: Inference answers are almost never the most extreme option. Watch for words like 'completely,' 'always,' or 'universally' — they often mark the trap.

Q6. The passage suggests that DeepSeek's emergence was significant primarily because

Answer: A

DeepSeek built 'a powerful product at a lower cost,' which is what put distillation on the map. C is tempting but wrong — DeepSeek didn't admit anything; OpenAI made the accusation (Trap B — uses passage vocabulary in wrong combination). SAT Tip: Watch carefully for who said or did what — passages often pair real names with real actions, and distractors swap them.

Q7. The author's tone in describing the White House accusations is best described as

Answer: B

The article presents Kratsios's claims, then quotes the Chinese embassy calling them 'pure slander,' then includes outside experts — classic balanced reporting. D overstates the emotional register (Trap C — true that tensions are real, but the author isn't 'alarmed'). SAT Tip: Tone questions reward you for noticing what the author chose NOT to say. A neutral reporter quotes both sides; an advocate cherry-picks one.

Q8. The author's primary purpose in this article is to

Answer: B

The article informs readers about a specific policy escalation, providing context, quotes, and consequences. A states a policy position the author doesn't take (Trap C — McGuire suggests bans, but the author reports his view, not endorses it). SAT Tip: Distinguish what sources in an article argue from what the author argues. A reporter quoting an opinion isn't the same as the reporter holding that opinion.

Q9. Which statement about the relationship between computing power and distillation can most reasonably be inferred?

Answer: C

The expert quoted says distillation helps 'offset deficits in AI computing power,' implying it's a workaround when chips are scarce. A is the opposite (Trap A — wrong direction). SAT Tip: When a passage describes a problem and a tactic, the tactic is usually a partial solution to the problem — not unrelated to it and not the cause of it.

Q10. Which choice provides the best evidence for the answer to the previous question?

Answer: B

B directly states that distillation 'offsets deficits in AI computing power' — the exact mechanism the previous answer described. C is true but addresses model quality, not the computing-power workaround (Trap C — passage-true but not relevant to the specific inference). SAT Tip: On evidence-pairing questions, find the line that PROVES your previous answer — don't pick a quote just because it's accurate, pick the one that's logically necessary.