Anthropic, AI, and the US Department of War
Due March 12 by 11:59PM

Policy Memo Assignment
This assignment is drawn from current headlines. As a result, events over the next two weeks will likely inform the recommendations you make in your memo.

Form of group of three or four students. Smaller or larger groups are not permitted. In our experience, interdisciplinary teams tend to produce stronger memos.

Each team should submit one copy of the final memo as a PDF file. The length of the memo should be the same whether your team is composed of 3 or 4 individuals. Your memo should be no more than 3,000 words (excluding references).
Late policy:

- Submitted within 3 days of deadline: one grade penalty per day
- Submitted more than 3 days after the deadline: no credit

After the assignment is submitted, we will ask students to complete a short survey on-line that enables you to communicate how the workload was distributed across team members. If the distribution of the workload is not equal, we reserve the right to increase or decrease grades for individual team members accordingly.

Statement on the Use of Automated Writing/Coding Tools
Generative AI provides opportunities to support learning but also opportunities to cheat. The use of automated writing or coding tools (e.g., ChatGPT, Gemini, or Claude) for submitting assignments is permitted (though not required) in this class, consistent with Stanford's policies on plagiarism. All students must submit (as an appendix to their paper) a document that discloses their use of AI tools. This appendix should include:

- List of the AI tools or systems used by the student (e.g., Claude, ChatGPT, Gemini, etc.)
- Complete and fully inclusive catalogue of the student-generated prompts

Stanford's plagiarism policy:
https://communitystandards.stanford.edu/policies-and-guidance/bja-guidance-definitions-and-clarifications/what-plagiarism/plagiarism

Background
On February 26, 2026, Anthropic published a public statement regarding its contract with the Department of War (formerly the Department of Defense). In the statement, CEO Dario Amodei described Anthropic as "the first frontier AI company" to deploy models in classified U.S. government networks and at National Laboratories, and noted that Claude is used extensively across defense and intelligence agencies for analysis, simulation, and cyber operations. The statement identified two applications that Anthropic would not support regardless of government pressure: the use of AI for *mass domestic surveillance* of U.S. persons, and the deployment of *fully autonomous weapons* that remove human decision-making from targeting. Amodei

described these as “bright red lines” that the company could not “in good conscience” cross, while expressing willingness to facilitate a smooth transition to alternative providers if the government chose to end its relationship with Anthropic.

The statement was issued in the wake of an escalating confrontation between Anthropic and Secretary of War Pete Hegseth. The Department of War had demanded that all AI contractors accept “any lawful use” of their models and remove internal safety restrictions. When Anthropic refused, the Pentagon threatened to strip the company of its $200 million contract, designate it a “supply chain risk”—a classification ordinarily reserved for companies doing business with adversarial nations such as China—and, most dramatically, to invoke the Defense Production Act to compel Anthropic to provide its models for military use regardless of the company’s objections.

The $200 million contract was awarded to Anthropic in July 2025 by the Pentagon’s Chief Digital and Artificial Intelligence Office (CDAO) as a two-year prototype agreement. This deal was part of a broader $800 million initiative that granted similar $200 million ceilings to Google, OpenAI, and xAI to develop "agentic" AI capabilities for mission-critical tasks like intelligence analysis and cyber operations. Anthropic's agreement focused on deploying its Claude Gov models within classified networks

The Defense Production Act and “Supply Chain Risk”: Brief Explainers

The Defense Production Act (DPA) is a Korean War–era statute, originally signed by President Truman in 1950, that grants the executive branch broad authority to direct private industry in the name of national defense. Its core provision—Title I—allows the president to designate specific goods or capabilities as “critical and strategic” and to require private businesses to accept and prioritize government contracts for those goods. Title III authorizes the government to make loans, purchase materials directly, and repurpose production facilities to expand capacity. Title VII provides for voluntary agreements and advisory committees.

The DPA has been invoked in a wide range of contexts since 1950: to manage defense production during the Korean War and Cold War, to ensure energy supplies during the California energy crisis, to compel General Motors to manufacture ventilators during the COVID-19 pandemic, and to accelerate baby formula production during the 2022 shortage. President Biden also invoked the DPA in a 2023 executive order on AI, requiring companies to share safety test results with the government—an order that President Trump subsequently repealed.

However, as legal scholars at Lawfare and elsewhere have noted, the DPA has never been used to compel a company to produce a product that it has deemed unsafe, or to override a company’s terms of service. Secretary Hegseth’s threat to invoke Title I against Anthropic would represent an unprecedented use of the statute—one that would raise significant questions about the limits of executive authority over private technology firms and the extent to which the government can override corporate safety commitments in the name of national defense.

The U.S. federal government's authority to designate a company a "supply chain risk" derives primarily from the Federal Acquisition Supply Chain Security Act of 2018 (FASCSA), which established the Federal Acquisition Security Council (FASC) — an interagency body co-chaired by the Office of Management and Budget, the Department of Homeland Security, and the Director of National Intelligence. The FASC is empowered to recommend exclusion and

removal orders against companies whose products or services are deemed to pose supply chain risks to the federal government. The designation was designed with foreign adversaries in mind: its most prominent use has been against companies like Huawei and Kaspersky Lab, where concerns about ties to the Chinese and Russian governments, respectively, led to government-wide bans on the use of their products. Once a company is designated a supply chain risk, federal agencies and their contractors are generally prohibited from procuring its products or services, effectively cutting the company off from the entire federal contracting ecosystem.

What makes the threat against Anthropic unusual — and, in the view of many legal scholars, unprecedented — is that the designation is being wielded not against a foreign-linked firm suspected of espionage or sabotage, but against a domestic American company in what is essentially a policy disagreement over the terms of use for its AI models. The "supply chain risk" label carries enormous practical consequences: if applied to Anthropic, it would not only terminate the company's own government contracts but would also require every defense contractor and federal vendor to certify that they do not use Anthropic's technology, effectively blacklisting the company across the entire defense industrial base. Critics argue that repurposing a national security tool designed to counter foreign threats as leverage against a domestic company's safety policies represents a significant expansion of executive power — one that could chill corporate governance commitments across the technology sector.

Geopolitical and Competitive Context

The confrontation between Anthropic and the Department of War is unfolding against a backdrop of heightened geopolitical tension. As of late February 2026, the United States is engaged in a significant military buildup in the Middle East amid an escalating crisis with Iran. The USS Gerald R. Ford and the Abraham Lincoln carrier strike group are deployed to the Arabian Sea, and senior administration officials have publicly stated that there is a high probability of kinetic military action against Iran in the coming weeks. In this context, the administration’s demand for unrestricted access to frontier AI capabilities carries additional urgency—and additional risk.

At the same time, Anthropic faces intense competitive pressure from rivals that have adopted a more accommodating posture toward military use. OpenAI, Google, and Elon Musk’s xAI have all agreed to allow their AI tools to be used in “any lawful” military scenario without the restrictions Anthropic has maintained. xAI was recently approved for classified use. The Pentagon’s $200 million contracts were awarded to Anthropic, Google, OpenAI, and xAI last summer, and Anthropic was initially the preferred provider for classified applications—a position it now risks losing. Secretary Hegseth and other administration officials have publicly characterized Anthropic’s safety commitments as “woke AI,” a framing that carries political and reputational costs. If Anthropic is removed from defense networks, its competitors stand to absorb its share of military AI work, potentially weakening both the company’s business position and its influence over the norms of responsible AI development.

Your Task

Imagine you are writing a memo to **Dario Amodei, CEO of Anthropic, and the Long-Term Benefit Trust** that governs the company’s public benefit mission. The Long-Term Benefit Trust holds a controlling interest in Anthropic and is charged with ensuring that the company’s decisions serve the long-term benefit of humanity—a mandate that includes, but is not limited to,

AI safety. Your memo should advise Anthropic's leadership on how to navigate the current crisis: whether to maintain, modify, or reframe its "red lines," and how to position the company in a rapidly shifting landscape defined by government coercion, geopolitical volatility, and competitive pressure from rivals willing to operate without comparable restrictions.

Your memo should be no more than 3,000 words (excluding references). In your memo, address the following:

1. **Anthropic's stated position and the February 26 statement.** Evaluate the substance and strategic framing of Anthropic's public statement. Is the company's distinction between acceptable military uses (analysis, simulation, cyber operations) and its two "red lines" (mass domestic surveillance, fully autonomous weapons) a principled and defensible position? If these are lawful uses, according to United States law, on what grounds would or should a private company refuse such use?
2. **The Defense Production Act and Supply Chain Risk Designation threat.** Assess the legal and practical risks of the government invoking the DPA against Anthropic. How should the company prepare for this possibility? What legal strategies, public advocacy, or coalition-building could strengthen Anthropic's position? Consider the precedential implications: if the DPA can be used to override an AI company's safety policies, what does that mean for the industry as a whole? Also consider the Pentagon's threat to designate Anthropic a "supply chain risk" can be understood as an effort to shift the Overton Window on executive leverage over AI firms—normalizing coercive tactics that, if left unchallenged, may redefine what future administrations consider acceptable pressure on the private sector
3. **The geopolitical context.** How should Anthropic's leadership weigh the reality that the United States is on the brink of potential military conflict with Iran—and that AI capabilities are increasingly central to military planning and operations? Does the prospect of active hostilities change the calculus for the company's red lines, or does it reinforce the importance of maintaining them? How should Anthropic communicate its position in a wartime (or near-wartime) environment without appearing indifferent to national security?
4. **Competitive dynamics.** How should Anthropic respond to the fact that its major competitors—OpenAI, Google, and xAI—have agreed to allow unrestricted lawful military use of their models? If Anthropic loses its defense contracts and its competitors gain market share in national security applications, does that undermine or advance the cause of responsible AI development? Consider the argument that it is better to have a safety-conscious company at the table than to cede the space entirely to firms with fewer restrictions.
5. **Your policy recommendations.** Provide a concrete set of recommendations to Anthropic's CEO and Trust. The scope of these recommendations should be sensitive to unfolding events over the next two weeks. As a minimum, your recommendations should address (a) whether Anthropic should maintain, modify, or reframe its red lines; (b) how the company should respond to the potential (or actual) invocation of the DPA or "supply chain risk" designation; (c) how it should engage with the current geopolitical environment; (d) how it should position itself relative to competitors; and (e) any public

communications strategies that would strengthen the company's hand. Be specific and realistic: your recommendations should account for the possibility that Anthropic may lose its defense contracts regardless of what it does.

Interviews

Each team member must conduct **three interviews** with individuals who can provide expert or informed perspectives on the issues raised in this assignment. Your interviews should cover a range of viewpoints and must include representation from **at least three** of the following categories:

- **Employees at Anthropic or other major AI model developer companies.** This includes engineers, researchers, policy staff, or business development professionals who can speak to the internal dynamics of AI companies navigating government contracts, safety commitments, and commercial pressures. Interviews may be conducted on background if the interviewee requests confidentiality.
- **Veterans and members of the U.S. military.** This includes active-duty service members (to the extent permitted by regulation), veterans with recent operational experience, reservists, or military families. The goal is to understand how those who serve—or have served—view the integration of AI into military operations, the importance of restrictions on autonomous weapons and surveillance, and the obligations they believe technology companies owe to national defense.
- **National security and military policy experts.** This includes scholars at institutions such as the Hoover Institution, the Freeman Spogli Institute for International Studies (FSI) at Stanford, or comparable think tanks and research centers. These interviewees should be able to speak to the strategic dimensions of AI in military contexts, the legal and policy implications of the Defense Production Act, and the broader geopolitical competition between the U.S. and its adversaries.
- **Legal scholars and constitutional law experts.** This includes professors or practitioners specializing in executive power, administrative law, defense procurement, or the First Amendment—particularly those with expertise in the Defense Production Act or the legal rights of companies to refuse government contracts.
- **Civil society and AI safety and ethics advocates.** This includes representatives from organizations working on responsible AI, civil liberties, or the governance of emerging technologies—such as the ACLU, the Electronic Frontier Foundation, the AI Now Institute, or the Partnership on AI.

The voices of your interviewees should be included in the policy memo in some fashion. In an appendix (which does not count toward the word limit), provide a brief summary (150–250 words) that captures the interviewee's key arguments, areas of agreement or disagreement with your group's position, and any insights that shaped your recommendations. Include a methodology note describing how you selected your interviewees and what steps you took to ensure a diversity of perspectives.

Additional Directions

Format: Submit your memo as a single PDF document. Include a cover page with team members' names and the assignment option selected. Include the appendices in the pdf.

Citations: Use a consistent citation style throughout (e.g., Chicago, APA). All factual claims should be supported by citations to credible sources.

Suggested Resources
The following readings are recommended starting points. You are expected to conduct additional research beyond these sources.

On the Anthropic–Department of War Standoff:

- Anthropic (2026). "Statement on Department of War Discussions." *Anthropic Blog*, February 26, 2026.
- Hegseth's ultimatum and the supply chain risk designation: see coverage in *CNN Business* (Feb. 24, 2026); *Axios* (Feb. 24, 2026); *The Washington Post* (Feb. 24, 2026); *NPR* (Feb. 24, 2026).
- TechPolicy.Press (2026). "A Timeline of the Anthropic-Pentagon Dispute."

On the Defense Production Act:

- Lawfare (2026). "What the Defense Production Act Can and Can't Do to Anthropic."
- Congressional Research Service (2023). "The Defense Production Act of 1950: History, Authorities, and Considerations for Congress." CRS Report R43767.
- Council on Foreign Relations (2023). "What Is the Defense Production Act?"

On Responsible AI and Corporate Ethics:

- Anthropic's *Responsible Scaling Policy* (various versions)