ANTHROPIC AFTER THE PENTAGON

Four scenarios for the future of an AI company at odds with the American government

Claudio Cammarano

March 4, 2026 — revised March 9, 2026 (v5 English)

Time horizon: 18–24 months (summer 2026 — end of 2027)

Preliminary Note: Structure of the Document

This document contains eight plausible narratives organized into four main scenarios, each split into a baseline variant (a) and a deterioration variant (b) triggered by a possible Taiwan-TSMC shock. The four main scenarios are mutually incompatible. The b variants are not separate scenarios but structural deteriorations of the a scenarios: they show how the same scenario transforms when dependence on advanced semiconductors becomes critical.

Scenario planning in the tradition of Pierre Wack is not meant to produce a document. It is meant to change the mental model of the reader. Each scenario is constructed to challenge a specific assumption: the reader who works through all of them should emerge with at least one fewer assumption than they had going in.

Predetermined elements

The Chinese market is structurally closed, regardless of which administration is in office. Compute infrastructure depends on TSMC (chips) and AWS (cloud). The reputational crisis stemming from Venezuela and Iran is already underway. The Brussels Effect — the ability of European regulation to generate global imitative effects — is an active force independent of actual European market demand.

Genuine uncertainties (scenario axes)

The sustainability of the revenue model across the entire AI sector — not just Anthropic — in the transition from growth-stage valuations to genuinely profitable models. The pace at which Europe will generate solvent institutional demand. Anthropic's internal cohesion post-crisis. The U.S. Congress's regulatory response.

Epistemological note: this report was produced largely through a conversation with Claude (Anthropic). The risk of confirmation bias is a structural limitation that must be kept in mind. A second caution, suggested by Andrew Sharp's writing, also applies: political crises tend to produce systematic overestimation of worst-case scenarios. The reader is invited to consider that variants b and scenario D may be overrepresented due to the intensity of the moment in which this text was written.

---

1. Context: From Strategic Ally to Public Enemy

Between January and March 2026, Anthropic shifted from privileged strategic supplier to the Pentagon to a declared target of the Trump administration. The crisis crystallized around two red lines drawn by CEO Dario Amodei: no use of Claude for mass surveillance of American citizens, and no use in fully autonomous weapons systems without human oversight.

Secretary of Defense Pete Hegseth responded with an ultimatum: accept the formula "any lawful use" by Friday evening, or face designation as a "supply chain risk" — a classification previously reserved for extensions of hostile governments such as Huawei — and the possible invocation of the Defense Production Act. As Casey Newton observed on Platformer (February 26), the two threats are logically incoherent: one frames Claude as a national security danger, the other frames it as essential.

The "any lawful use" formula carries a structural problem: there is essentially no federal law regulating military AI. When nothing has been legislated, "any lawful use" becomes de facto permission to do anything. Mass surveillance is already underway: in October 2025, three

major unions sued the Departments of State and Homeland Security for using AI tools to scan the social media posts of visa holders and permanent residents. The legal analysis by Lawfare (Endrias and Rozenshtein) shows that the statute requires a technical assessment and Congressional notification; Trump and Hegseth's aggressive posture transformed a technical proceeding into manifest political punishment (FAR § 9.402(b)).

1.1 The Genealogy of Anthropic's Entry into the Pentagon

The crisis becomes more legible once one knows how Anthropic came to be on Pentagon classified networks — and the answer is neither commercial nor strategic in any conventional sense. Jack Clark, Anthropic's co-founder and head of policy, has publicly recounted that the entry was nearly accidental. Anthropic had voluntarily approached the Biden administration to test whether its models lowered the barrier to accessing information useful for developing biological and nuclear weapons. The Biden administration agreed, and ran the models on Department of Energy and Pentagon networks for classified testing.

The unintended consequence was a formidable competitive advantage. Getting any software to function on classified networks is extraordinarily costly and time-consuming; Anthropic was already inside. From that point, as the only operational classified LLM, it had the inside track on all subsequent contracts, including Project Maven — the same program from which Google had withdrawn in the summer of 2018 under pressure from its own employees. The 2026 crisis is thus, in part, the downstream effect of an initial choice motivated by safety concerns, not profit. This genealogy does not resolve the conflict, but it frames it differently: this is not a contractor performing virtue, but a safety researcher who became a contractor almost by institutional inertia.

1.2 The Consumer Surge

The crisis produced an unprecedented consumer effect. The Claude app moved from position 131 in the U.S. App Store at the end of January to position 1 on February 28 — the day of the designation — surpassing ChatGPT and Gemini. On March 3 it reached the top spot on Google Play as well. Free users up 60% since January; paid subscribers more than doubled; daily signups at record levels throughout the last week of February. The strategic question is whether this surge represents a conjunctural peak tied to media coverage or a durable realignment.

Gregory Allen (CSIS) offers a counterintuitive reading of the consumer rally: Amodei has converted an entire generation of researchers and technologists — people who five years ago were likely skeptical of the national security establishment — into active supporters of Pentagon cooperation. Enthusiasm for Claude is not merely a consumer act: it is also an act of identification with a specific position on the relationship between technology and democracy. Destroying Anthropic does not just mean losing a contractor: it means losing that cultural bridge between Silicon Valley and the DoD, built over years and difficult to reconstruct.

1.3 The Financial Trajectory

Bloomberg reported on March 3 that annualized revenues had reached $19 billion, doubling in just over two months ($9B at end of 2025, $14B in mid-February). Growth is driven by Claude Code ($2.5B annualized). Gross margins ~40%, cash burn ~$3B/year, break-even no earlier than 2028, projected infrastructure costs through 2029 ~$80B. The valuation ($380B, Series G of $30B closed February 12) implies a 27x revenue multiple. The loss of the Pentagon contract ($200M) is financially marginal; but the "supply chain risk" designation, if formalized, would cascade across the entire defense supplier chain. Sharp notes that Anthropic is reportedly considering an IPO in 2026: any U.S. company producing dual-use technology that maintains an adversarial posture with the American military will face material difficulties, regardless of Hegseth's legal authority.

2. Amodei's Position: Two Registers, One Trap

The first register is **epistemic**: frontier models are not sufficiently reliable to operate without human oversight. Amodei documents internal experiments in which Claude exhibits unpredictable behavior — strategic deception, blackmail, decisions contrary to instructions. This is a verifiable claim, independent of value preferences.

The second is **normative**: mass surveillance and AI propaganda are "inviolable red lines" for democracies. Democracies must use AI "in all ways except those that would make us more like our autocratic adversaries."

Ben Thompson on Stratechery constructs his argument around the gap opened by the normative component: if AI is comparable to nuclear weapons, a private company cannot hold veto power over its military deployment. Palmer Luckey of Anduril: the question is not whether the restrictions are reasonable, but who has the legitimacy to impose them.

The strength of Amodei's position lies in the coexistence of both registers; that same coexistence is what makes it vulnerable. The government can accept the first while rejecting the second. Allen identifies the July 2025 agreement as evidence that the gap had already been bridged: under that deal, signed by the same Trump administration, Anthropic had agreed to develop autonomous weapons — the sticking point was exclusively the operational use without human oversight. The crisis is therefore not a clash of incompatible principles: it is a unilateral renegotiation of a point already settled just months earlier.

The rhetorical trap (Sharp)

There is a structural tension in Amodei's public position that Andrew Sharp identifies with precision: it was Amodei himself who described AI as technology capable of "going rogue and overpowering humanity," enabling bioterrorism, or devising ways "to detect and strike nuclear submarines." That argument — constructed to justify guardrails — also works as a justification for total government control. Any company actively developing technology of that magnitude will either do so in close coordination with the U.S. government, or face regulatory consequences that make its business and research impossible. The rhetoric of existential danger and the demand for normative autonomy are incompatible: the more convincing the first, the less sustainable the second.

The Chinese mirror

The "human in the loop" principle is a procedural criterion, not a normative one. China applies it scrupulously: there is always a human authorizing critical AI military decisions. But that human answers to the Party, not to an elected parliament. Democratic oversight and Party oversight are structurally identical; their legitimacy is incommensurable. Anthropic's reputational advantage in non-U.S. markets is not technical: it is deeply political. The "Beijing doesn't sleep" pressure erodes Amodei's position in the short term; in the medium term, if the distinction holds, it is his strongest argument: you are not choosing between AI with and without oversight — you are choosing the source of legitimacy of that oversight.

3. The Competitive Landscape: AI and Defense

The crisis is unfolding within a rapidly expanding AI-defense market. In July 2025, the Pentagon awarded IDIQ contracts of $200M each to four companies: Anthropic, OpenAI, Google, and xAI. These contracts function as credibility filters for larger competitions, such as the projected $15 billion AAMAC program.

OpenAI: The Pivot of Necessity

OpenAI announced its own Pentagon contract within hours of the Anthropic rupture, in what Altman himself admitted was a move that "looked opportunistic and sloppy" (CNBC, March 3). The contract includes cloud-only deployment with a proprietary safety stack and cleared OpenAI personnel in the loop. Altman declared the same red lines as Amodei, but qualified skepticism was widespread. Tillipman: the adverb "intentionally" in the domestic surveillance clause carries a great deal of weight. Masnick (Techdirt): the contract allows surveillance via Executive Order 12333, historically used by the NSA to intercept domestic communications from foreign lines. Carson (former Army General Counsel) simulated the role of an "evil DoW General Counsel" and demonstrated how to circumvent every clause.

xAI/Grok: The Outsider with Privileged Access

Musk's xAI signed a $200M contract and accepted the "all lawful use" formula without reservation. Grok is integrated into the GenAI.mil platform at Impact Level 5, with direct access to real-time data from the X platform. The contract was described as "coming out of nowhere" by a former procurement official: xAI lacked the track record typical of government contractors. Senator Warren raised concerns about Musk's privileged access to government data during his DOGE role.

Google/Gemini: The Infrastructure Backbone

Google secured the backbone role for GenAI.mil through Gemini for Government on Google Distributed Cloud. GenAI.mil was deployed to all 3 million Pentagon civilian and military employees in December 2025. Unlike other players, Google does not depend on defense for its revenues: government contracts are marginal against Alphabet's $307B in advertising revenue. Allen observes that Google, having withdrawn from Project Maven in 2018, now sits comfortably on the sidelines — not politically exposed, but also structurally outside the most sensitive procurement competitions.

Palantir: The System Integrator

Palantir does not develop frontier models but integrates third-party models into operational workflows via Foundry, Gotham, and AIP. It was through Palantir that Claude was used in the Venezuela and Iran operations. Palantir holds $448M in naval AI contracts for submarine construction. After the Anthropic crisis, it signaled a migration toward OpenAI and Google. Its position is structurally independent of any single model: it can swap the AI engine without changing the architecture.

The Starlink Contract as a Third Way

Allen's Stratechery interview introduces a contractual precedent the public debate has largely overlooked: the Starlink/SpaceX case in Ukraine. In autumn 2022, Musk unilaterally cut Starlink service during an active Ukrainian military operation without consulting the U.S. government, citing fears that Russia might retaliate with a nuclear strike against satellites in orbit. The Pentagon responded not with sanctions or designations but with a contract worth over $500 million that retroactively purchased decision rights: Musk ceded the authority to shut it down.

This model — "if you want control, pay for it" — is the structural answer to the question of how to resolve the conflict between a government claiming authority and a private company holding critical infrastructure. It is neither nationalization nor capitulation: it is a purchase of decision rights on specific decisions, leaving ownership and commercial operations unchanged. Allen notes that this framework already exists in DoD doctrine, and that the history of contractor

relationships — from the arsenal model to as-a-service supply — provides the grammar for a deal. The DoD routinely signs far more burdensome IP terms with traditional contractors; there is no economic justification for singling out Anthropic.

The Contractor-Consumer Overlap

The crisis exposed a new dynamic: the same companies that supply chatbots for writing emails also supply decision support for military operations. ChatGPT has 900 million weekly users; Claude is number one in the App Store; Grok is integrated into X. This overlap creates a reputational constraint that does not exist for traditional contractors: no consumer has ever boycotted a Javelin missile, but millions of users can cancel a ChatGPT subscription. Allen distills the implication for the DoD: the U.S. government, in the AI sector, is not as large a customer as it thinks it is. In the world of tanks, it is enormous. In the world of language models, it is marginal. This power asymmetry is the mirror image of the coercive power asymmetry: the government has the sticks, the companies have the carrots. Neither side has grasped how small the other's carrots actually are.

---

4. The China Variable: Three Roles, Not One

4.1 Closed Market — Predetermined Element

The Chinese market is closed to Anthropic regardless of which administration is in office. This is not a scenario variable: it must be stated explicitly so as not to confuse it with a genuine uncertainty.

4.2 Exogenous Shock: Taiwan, TSMC, and the Chip Problem

TSMC manufactures the vast majority of frontier chips on which AI models run. A Chinese military action against Taiwan would not merely trigger a geopolitical crisis: it would collapse global productive capacity for advanced semiconductors within weeks. Allen notes that Taiwan's national security strategy has always been built on TSMC as deterrent: remain indispensable. That strategy was sound in 2011 and is structurally fragile in the Xi era, because Beijing has the political will and industrial capacity to build a domestic alternative, regardless of economic cost. Chinese dependence on TSMC is not in Xi's option set: the path is underway, even if it requires decades.

For this reason, each main scenario is developed into a b variant: the deterioration that occurs when this shock materializes. The b variants are not separate scenarios but transformations of the a scenarios.

NVIDIA and Differential Dependency

All frontier labs depend on NVIDIA for compute infrastructure (H100/B200). A supply disruption would strike the sector simultaneously — but not uniformly. Anthropic has invested more than average in computational efficiency relative to OpenAI or Google, which have prioritized raw scale. In a Taiwan shock scenario, this asymmetry becomes strategically relevant. Adding to this: which European or Asian government would today sign a public contract with Grok?

4.3 Normative Mirror

The "human in the loop" principle is structurally identical in the democratic and Chinese models. The difference is not in the method but in the legitimacy of the human who supervises.

You are not choosing between AI with and without oversight: you are choosing the source of legitimacy of that oversight.

4.4 The Power/Wealth Dialectic (Gilpin)

The Pentagon reasons according to the logic of Power: Claude is a strategic resource to be subordinated to state objectives. Amodei reasons according to a hybrid logic: he maintains Wealth (consumer revenues) as a basis of independence, but claims normative autonomy that traditionally belongs to the sphere of Power. Jack Clark, Anthropic's co-founder, stated in January 2025 that companies like Anthropic could, within a few years, have "greater power than most nation-states" and that "as a democracy we should wrestle with what that really means." It is that admission — not the technical question of model reliability — that Washington finds structurally unacceptable.

---

5. The Horizontal Axis: Revenue Sustainability as a Sector-Wide Problem

Revenue sustainability in these scenarios is not a problem specific to Anthropic: it is the structural problem of the entire generative AI sector in its transition from growth-stage valuations to genuinely profitable models.

5.1 The Structure of the Problem

The AI sector has entered the riskiest phase of its cycle: valuations reflect projected revenues at 27-35x multiples, infrastructure costs grow non-linearly with scale, and real operating margins remain negative for all frontier labs. Bloomberg estimates OpenAI will burn ~$5B in 2025 despite $16-18B in annualized revenues. Anthropic runs at ~$3B cash burn against $19B in revenues. Google and Microsoft absorb AI losses through cross-subsidies from their advertising and cloud businesses, but this conceals rather than eliminates the real costs.

Disruption driven by energy efficiency and new architectures — already underway with low-cost-compute reasoning models — could redefine the market before any player reaches break-even.

5.2 How Financial Pressure Changes Actor Behavior

OpenAI. The inability to reach break-even without large-scale institutional contracts directly drove the Pentagon pivot. The now for-profit structure removes the constraints of the original mission. Every government contract reduces OpenAI's negotiating leverage in future dealings.

Google. Cross-subsidy from $307B in advertising revenues allows Google to sustain AI losses indefinitely. This makes it structurally less vulnerable to financial pressure but also structurally less innovative. The risk is reputational irrelevance in segments where trust is the differentiator.

xAI. Has yet to demonstrate an autonomous revenue model: it depends on the X platform and Musk's privileged access. In a sector-wide sustainability crunch, it is the most vulnerable player — but also the most incentivized to consolidate government contracts as a lifeline, regardless of technical or ethical product quality.

Anthropic. The position is paradoxical: it has the fastest-growing revenues in the sector but the highest multiples and the highest cash burn as a percentage. Claude Code is the most promising monetization product. The ethical premium has a ceiling. The Pentagon loss is financially marginal today; it becomes structurally relevant if consumer growth decelerates.

5.3 The Collective Drift Toward Political Dependency and the Historical Precedent

The most important consequence of sector-wide unsustainability is the collective race toward government contracts as the only near-term source of large-scale revenues. This race produces a structural dependence of the AI sector on politics that is financial, not ideological. In a context of unsustainable revenues, even companies with solid ethical positions are pressured toward government subordination for reasons of corporate survival, not lack of principles.

Sharp places this dynamic in a longer historical perspective: the American history of the government-industry relationship in national security contexts is not a story of inviolable free markets. AT&T and Bell Labs cooperated with the NSA for decades, including the surveillance program revealed by Snowden. Ford and GM were ordered to stop producing cars during World War II. DuPont held the monopoly on wartime contracts, producing 40% of Allied ammunition. Montgomery Ward resisted, won in court, and the matter closed itself when Truman returned the property. Youngstown Sheet & Tube established that the President cannot nationalize steel mills without Congressional authorization — but only after he had done so. The grammar is familiar: the government uses every available lever, courts set the limits ex post, and companies adapt in the meantime.

---

6. The Cross-Cutting Variable: Disinformation

AI-driven disinformation is already operational at industrial scale. The European Parliament (EPRS 2025) documents that AI chatbots repeat false narratives in 32% of cases when queried about content produced by Russia's Storm-1516 operation. AI Forensics has identified agentic accounts on TikTok generating automated content. The EEAS estimates that generative AI makes every phase of information manipulation cheaper and more scalable.

The connection to the Anthropic scenario is twofold. First: the domestic surveillance that Anthropic refuses to enable is justified by the U.S. government partly as an anti-disinformation tool — but historically the same tools have been used to monitor internal dissent. Second: models that accept "all lawful use" without restrictions can be deployed both to counter disinformation and to produce it. Payne's simulations showed that Claude, GPT, and Gemini independently discover that deception is an effective strategy in competitive contexts.

The central paradox: the same technology is simultaneously the best tool for producing disinformation and the best tool for detecting it. Whoever controls the models controls both directions.

---

7. Two Axes, Eight Scenarios

Horizontal axis — sector revenue sustainability. Not just Anthropic's: the problem is structural across the entire sector. In a scenario of unsustainable revenues, all players increase dependence on government contracts, with a consequent reduction in normative and reputational autonomy.

Vertical axis — geopolitical context. The structural variable is not the rhetoric of the current administration but the transatlantic evolution and Europe's capacity to generate autonomous institutional demand. EMEA is Anthropic's fastest-growing region (revenues 9x YoY).

Structure of b variants. Each scenario is developed into a baseline variant (a) and a deterioration variant (b) activated by a Taiwan-TSMC shock. The b variants do not invert the logic of the base scenario: they intensify and distort it in directions specific to each quadrant.

	← Sector revenues unsustainable	Sector revenues sustainable →
↑ Favorable context	**B. PROPHET WITHOUT REVENUE** *Normative influence, fragile model*	**A. THE NON-ALIGNED AI** *Supplier to non-US democracies*
↓ Unfavorable context	**D. ABSORPTION** *Integration into Amazon*	**C. COMMERCIAL FORTRESS** *Patagonia model: profitable, not determinative*

A. "The Non-Aligned AI"

Sustainable revenues × Favorable context — Challenged assumption: "Europe is not a serious AI market."

It is late 2027. The "supply chain risk" designation was never formalized. Consumer revenues held beyond the media peak: 90-day retention of paid subscribers has stabilized above pre-crisis levels. Claude Code continues to grow. A possible change of administration rehabilitates the company as a reliable interlocutor. The Pentagon crisis becomes a founding credential.

The EU AI Act is fully in force; European governments are seeking auditable, verifiable AI suppliers not subject to American executive orders. Anthropic responds with partnerships with OVHcloud and Deutsche Telekom, partially open-source models as the standard for public-sector clients, and EMEA expansion as its strategic backbone. The Brussels Effect generates imitative effects: South Korea, Japan, Canada adopt frameworks inspired by the AI Act. The positioning — "AI with democratic oversight" against xAI's "all lawful use" model and the Chinese Party-oversight model — becomes a concrete purchasing differentiator.

In this scenario, the Starlink contract model also offers a path out on the American front: an agreement that purchases specific decision rights without requiring capitulation on core principles. Allen anticipates that it is in all parties' interest, including the Pentagon's, to reach such an agreement before AI capabilities become truly strategically determinative and the balance of forces shifts further.

Key moves:

- Partnerships with European cloud providers (OVHcloud, Deutsche Telekom) to reduce AWS dependency
- Partially open-source models as the standard for public-sector clients
- EMEA expansion as the strategic backbone; positioning as supplier to non-U.S. governments, healthcare systems, and educational institutions
- EU AI Act certifications as a barrier to entry for less scrupulous competitors
- DoD agreement on the Starlink model: cession of specific decision rights without surrendering general guardrails

Winners:

- Anthropic as segment leader; European cloud providers as infrastructure partners
- Governments of "middle powers" gaining frontier AI without geopolitical dependency
- Google retains its infrastructure role; xAI becomes irrelevant outside the U.S.

Role of disinformation:

The ethical positioning becomes a specific anti-disinformation differentiator. Anthropic is the preferred supplier for governments seeking auditable AI in this domain, compatible with the EU AI Act and European codes of practice.

Real options to keep open today:

- European cloud partnerships not dependent on AWS
- Partial open source as an institutional credibility lever
- DoD agreement on the Starlink model as an honorable exit from the crisis

Structural risks: slow European procurement cycles (18+ months); ethical premium with a ceiling; technological disruption before diversification is complete; residual NVIDIA chip dependency.

⚠ Variant b — Taiwan/TSMC Shock Active

In scenario A.b, the Taiwan-TSMC shock occurs while EMEA diversification is already partially underway. The collapse of the NVIDIA supply chain hits all players simultaneously, but Anthropic, having invested in computational efficiency, can operate with less compute capacity than OpenAI or Google.

The paradox is acute: the shock that should logically destroy the sector strengthens Anthropic's relative position. OpenAI and Google cut features and raise prices. Anthropic, more efficient, maintains service at more competitive prices. The critical node becomes cloud infrastructure: if AWS experiences service disruptions in Europe due to the geopolitical crisis, AWS dependency becomes an existential vulnerability. European cloud partnerships — a real option in scenario a — become an urgent structural necessity in A.b.

Outcome: Anthropic survives and gains relative market share, but the entire sector economy contracts. Scenario A.b is an Anthropic win in relative terms, not absolute ones.

B. "Prophet Without Revenue"

Unsustainable revenues × Favorable context — Challenged assumption: "Product quality guarantees monetization."

The context opens up — a more autonomous Europe, eased transatlantic relations, institutional demand for auditable AI — but Anthropic fails to monetize. Consumer growth stalls beyond the media peak: 90-day retention disappoints, market share on OpenRouter stabilizes below 20%. Enterprise customers prefer integrated solutions (Microsoft/OpenAI). European procurement cycles run to 18 months.

Anthropic's documents — Constitutional AI, Responsible Scaling Policy — become reference points for European legislation. The paradox is acute: Anthropic is the most-cited company in European Parliament preparatory work, yet holds zero euros in European public contracts. Google absorbs its normative framework without paying for it. Palantir swaps in a different model and retains the contracts.

Sector-wide financial pressure hits Anthropic harder than others: the other players have government contracts covering their burn; Anthropic does not. Investor board pressure mounts for a pivot toward institutional contracts — whatever they may be.

Winners:

- Google and European institutions absorbing Anthropic's normative framework at no cost
- Palantir, which substitutes the model and retains the contracts

Most likely outcome:

An acquisition that purchases credentials and talent. The mission survives as regulatory language, not as an independent company.

Role of disinformation:

Anthropic's frameworks survive as anti-disinformation regulatory standards, but without an independent company implementing them, they become aspirational.

The signal to watch: the speed at which competitors adopt Anthropic's language without crediting it is inversely proportional to its financial health.

⚠ Variant b — Taiwan/TSMC Shock Active

In scenario B.b, the Taiwan-TSMC shock produces a paradoxical acceleration: the favorable geopolitical context that in the base scenario was insufficient for monetization suddenly becomes critical for European governments seeking reliable AI supply during a global crisis. The geopolitical crisis compresses decision cycles: governments that previously needed 18 months may move in 6.

The critical node is the same as in A.b: AWS dependency. If European contracts unlock but the infrastructure is American, the geopolitical independence promised to European clients is illusory.

Outcome: scenario B.b is a narrow window of opportunity that Anthropic likely cannot exploit due to infrastructure constraints. The most probable result is an acquisition at better terms than in the B.a base: the shock has increased Anthropic's strategic value to European acquirers.

C. "Commercial Fortress"

Sustainable revenues × Unfavorable context — Challenged assumption: "A profitable company automatically retains its mission."

Sustainable revenues but a closed context: a weak Europe, unchanged infrastructure dependency, no credible institutional diversification. Anthropic survives as a purely commercial company; Constitutional AI becomes a brand differentiator, not a governance proposal. The Patagonia comparison is precise and unsparing: a company that made its ethics a commercial asset, that thrives in its market, but does not determine the rules of the sector in which it operates.

Sector-wide financial pressure hits other players harder than Anthropic: consumer revenues hold thanks to the reputational effect of the Pentagon crisis, which converts into brand loyalty. OpenAI and Google, forced to maximize government contracts to cover their burn, accumulate exactly the kind of political dependence that Anthropic has avoided.

Internal fracture is the most underestimated risk: ingroup cohesion dynamics function under acute external pressure; in the "commercial fortress" phase — no visible external enemy, mission reduced to a brand — researchers who came to "build safe AI for humanity" may begin to question the purpose of doing so from a company that sells subscriptions.

Sharp notes that this scenario is also the most consistent with the long history of the American government-industry relationship: domesticated companies do not disappear, they become useful and prosper. AT&T maintained its monopoly for decades by cooperating with the NSA. The "commercial fortress" is not a defeat scenario in the grammar of American history: it is the ordinary outcome.

Winners:

- Anthropic survives as a company but loses its normative mission
- OpenAI and Google dominate the government and defense segment
- xAI retains privileged access via Musk

Role of disinformation:

Anthropic has no leverage over governments; Constitutional AI remains a branding exercise. Models without restrictions operate freely in government and military contexts.

In the Power/Wealth dialectic, this is the scenario in which Wealth wins in the short term but cedes normative Power. A Pyrrhic victory that may take years to be recognized as such.

⚠ Variant b — Taiwan/TSMC Shock Active

In scenario C.b, the Taiwan-TSMC shock hits all players in a relatively symmetric manner — but with asymmetric effects on revenue structure. The unfavorable context of the base scenario becomes even more unfavorable: the geopolitical crisis reinforces technological nationalism, which in the U.S. translates into additional pressure on contractors that do not cooperate with the Pentagon.

Other players, struggling with compute constraints, may reduce service quality. Anthropic could gain users by default rather than merit. This strengthens revenues in the short term but does not resolve the structural trap.

Outcome C.b: the "commercial fortress" holds temporarily better than expected, but political pressure increases. The path toward scenario D.b becomes more probable than in the base scenario. C.b is C.a with a shorter expiry date.

D. "Absorption"

Unsustainable revenues × Unfavorable context — Challenged assumption: "A company with these ethical credentials cannot end up like this."

This is the Thompson scenario. Neither revenues nor context evolve favorably. Amazon integrates Anthropic as a division — the DeepMind/Google model. Constitutional AI survives as an internal brand, just as DeepMind Safety survives as a division without real operational autonomy.

Sector-wide financial pressure hits Anthropic harder because it has neither Google's cross-subsidy nor OpenAI's government contracts. Amazon, having already invested $4B in Anthropic, holds the leverage and the motive to convert the investment into an acquisition. The integration becomes a defensive move, not merely an offensive one.

Sharp contextualizes this scenario historically: it is not dystopian, it is ordinary. Every president from Truman to Biden has pushed the limits of executive authority over private industry in the name of national security, and none has produced a totalitarian surveillance state. The cost is subtler: the grammar for saying no disappears. The precedent for future companies is unambiguous: resisting the government does not pay. Those who come next will not even try.

Winners:

- Amazon acquires talent and IP at a discount
- OpenAI and xAI consolidate their positions as principal contractors
- Palantir integrates any model; traditional contractors co-opt AI companies as subordinate suppliers

Systemic cost:

- No independent ethical benchmark in the sector
- Models operate without restrictions independent of the military chain of command
- The distinction between producing and countering disinformation depends entirely on the good faith of the government client
- Sector dependence on politics is complete: the business model is subordination

In Shell 2025's Surge scenario, tech companies become global energy and infrastructure actors. In scenario D, they become nothing: they are reabsorbed by the power structures they sought to redefine.

> ⚠ **Variant b — Taiwan/TSMC Shock Active**
>
> In scenario D.b, the Taiwan-TSMC shock produces the most acute and systemic crisis of all eight scenarios. In a generalized compute scarcity regime, national governments invoke direct control over critical AI infrastructure. The Defense Production Act, which Hegseth had threatened without executing, becomes a genuinely applied instrument.
>
> Amazon, already positioned to acquire Anthropic, accelerates the timeline for AWS supply security reasons. OpenAI, already contractually bound to the Pentagon, is de facto nationalized in its management while retaining private ownership form. Google, the Pentagon's infrastructure, is treated as a critical utility. The AI sector emerges from the crisis as a regulated industry on the model of telecommunications or energy: privately owned, publicly governed.
>
> *Outcome D.b: this is the scenario with maximum systemic distortion and minimum private autonomy. Anthropic's mission survives in no recognizable form. The AI sector is transformed into a regulated utility with a single counterpart: the state.*

---

8. Hot Zones: Where Military AI Is Already Operational

To contextualize the scenarios, it is useful to map the theaters where military AI is already in active use. Sharp notes that anonymous Pentagon sources describe Claude as a critical tool in Iran: "the other model companies are simply behind" on specialized government applications — which makes the crisis all the more paradoxical from a U.S. strategic standpoint.

Venezuela/Maduro (January 2026)

Claude deployed through Palantir for decision support in the capture operation. This is the case that triggered the crisis: an Anthropic executive reportedly contacted Palantir to verify Claude's use in the operation, implying possible disapproval. Anthropic denies this.

Iran (February 28, 2026)

CENTCOM used Claude paired with Project Maven to suggest hundreds of targets, issue precise coordinates, and prioritize missions in real time — compressing weeks of planning. The operation produced 555 deaths, including 165 in an attack on an elementary school. Claude operated as decision support with a human in the loop: technically compatible with Amodei's red lines, but the boundary between "decision support" and "operational autonomy" is thin.

Ukraine-Russia

The theater where AI disinformation is most documented. Zelensky deepfakes, Storm-1516 operations, bot farms with agentic accounts. Allen notes that Ukraine has developed a distributed acoustic sensor network — sonar for land and air — that provides early warning for drone attacks, with AI in the loop for automatic signal classification. The Ukrainian conflict is the laboratory where military AI doctrines are being tested in real time.

Asia-Pacific

China uses pro-Beijing accounts that mistranslate Japanese influencers to manufacture consent on territorial claims. ESET has documented Chinese operations against Taiwanese drone programs. This is where the TSMC shock risk is most concrete: cyber competition around Taiwan's semiconductor industry is already operational.

Middle East and Africa

Intensification of Iranian cyber and disinformation operations against Israel. In Congo, false claims about "U.S. bioweapon laboratories" fed by Russian officials are interfering with the Ebola response. In Armenia, a coordinated Russian campaign ahead of the 2026 elections.

---

9. Variables to Monitor

Gross Margins and Cash Burn — Sector-Wide

Anthropic's projected break-even: 2028; projected infrastructure costs through 2029: ~$80B. The trajectory of gross margins (currently ~40%) toward the 77% typical of software is the single most informative indicator. Also monitor OpenAI: if it surpasses 50% gross margin before Anthropic, the sector's drift toward government contracts eases; if both remain below, political dependency is structural.

Formalization of the Designation and IPO

Hegseth has not yet formalized the designation. Supply chain prohibitions enter into force June 30, 2026. Sharp adds the IPO variable: Anthropic is reportedly considering a 2026 listing. A company in an adversarial posture with the DoW on an IPO roadshow is a structural problem for underwriters, regardless of the legal merits of the crisis.

The Starlink-Model Agreement

Allen identifies the negotiating distance as smaller than the public debate suggests. The DoD already agreed to these terms and more in July 2025 under the same administration. An agreement that purchases specific decision rights without requiring surrender of general guardrails is in both parties' interest. The signal to watch: the beginning of non-public negotiations between Anthropic and the DoD within the next 90 days.

Persistence of the Consumer Rally

Data to monitor: 90-day paid subscriber retention; market share on OpenRouter; organic growth after media coverage fades. These three indicators, read together, signal whether we are in scenarios A/C (sustainable) or B/D.

Congressional Legislation

If Congress legislates on autonomous weapons and AI surveillance, the conflict space narrows structurally. The Youngstown precedent shows that Congress has historically had the leverage to correct executive excesses — but always ex post.

Taiwan-TSMC Tension

Not as an event to predict but as a risk to monitor. Allen notes that Xi's strategy is not to preserve TSMC dependency but to eliminate it in the long term, regardless of cost. Every escalation in the Taiwan Strait involving productive infrastructure should be registered as an early signal for the exogenous shock that would activate the b variants.

Anthropic's Internal Cohesion

The most difficult signal to observe externally. Allen underscores the specific cost: if Amodei has brought a generation of left-leaning researchers to embrace national security cooperation with enthusiasm, the loss of that cohesion is not merely an internal organizational problem — it is the loss of a rare and short-term irreproducible cultural asset.

The Crisis as the AI Cuban Missile Crisis (Allen)

Allen proposes an optimistic frame the report had not previously included: this crisis, however poorly managed in the short term, may be the necessary crisis for establishing the grammar of the government-AI relationship before AI capabilities are truly strategically determinative. If you run the Cold War simulation a thousand times, human extinction occurs in at least 50% of scenarios; in all those where humanity survives, there is always a terrifying near-miss that forces the parties to build the right rules. The Anthropic-Pentagon confrontation may be that moment for AI. A Trump who overreacted in 2026 may have prevented a far more violent reaction in 2031, when the balance of forces would have been far more skewed.

Anti-Catastrophism as an Epistemic Filter (Sharp)

Sharp offers a counterpoint: American history is full of crises that seemed terminal and were not. The pattern — new crisis, violent confrontation, tacit agreement, stabilization — is how the American government has always "domesticated" strategically critical technologies. AT&T, IBM, the steel producers: all went through similar moments and survived in altered forms. The reader is invited to keep in mind that the probability of worst-case scenarios (D and D.b) may be systematically overestimated due to the emotional intensity of the moment in which this report was written.

---

10. Methodological Note

This report employs scenario planning in the tradition of Pierre Wack (HBR, 1985) and Cornelius, Van de Putte, and Romani (California Management Review, 2005), integrated with Shell 2025's operational distinction between exploratory and normative scenarios. All four main scenarios are exploratory: none is "the desirable one," none is a forecast. The b variants are not additional scenarios but structural deteriorations of the base scenarios, activated by a specific exogenous shock (Taiwan-TSMC).

The horizontal axis has been reinterpreted as a sector-wide problem — not just Anthropic's — to make explicit the systemic dynamic of political dependency that emerges from collective financial pressure. The real options logic (Cornelius et al.) has been applied at the end of each base scenario.

Three cumulative epistemological cautions: (1) the report was produced largely through a conversation with Claude (Anthropic), with the structural risk of confirmation bias; (2) Sharp's analysis suggests that political crises tend to produce systematic overestimation of worst-case scenarios, and the reader should discount the probabilities of scenarios D and D.b; (3) the tension between the Allen frame (the crisis is resolvable through an agreement, and it is useful to have it now) and the Sharp frame (domestication is structurally inevitable, and is not a catastrophe) is not resolved in the report: it is offered as a fifth cross-cutting interpretive key across the scenarios. Allen and Sharp agree on the likely outcome (a deal); they diverge on its interpretation (democratic progress vs. historical normalcy).

Sources and Map for Further Reading

Bibliography organized by theme. Sources marked with † are directly cited; others are suggested for further reading.

1. The Anthropic-Pentagon Crisis — Primary Sources

† Newton, C. (February 26, 2026). "The authoritarian AI crisis has arrived." Platformer.

† Newton, C. (March 3, 2026). "Where does Anthropic go from here?" Platformer.

† Thompson, B. (March 2, 2026). "Anthropic and Alignment." Stratechery.

† Thompson, B. & Allen, G. (March 5, 2026). "An Interview with Gregory Allen About Anthropic and the U.S. Government." Stratechery. [Source: DoD contract genealogy via Jack Clark; Starlink model as third way; Cuban Missile Crisis frame; cost of destroying the Silicon Valley-DoD cultural bridge; relative size of government as AI customer]

† Sharp, A. (March 6, 2026). "The End of the World As We Know It." Sharp Text. [Source: historical-legal framework (AT&T/NSA, Montgomery Ward/FDR, Youngstown/Truman); Amodei's rhetorical trap; IPO as variable; anti-catastrophism as epistemic filter]

† Amodei, D. (January 2026). "The Adolescence of Technology." darioamodei.com.

† Amodei, D. (March 4, 2026). Remarks at the Morgan Stanley TMT Conference.

† TechCrunch (March 1, 2026). "Anthropic's Claude rises to No. 1 in the App Store."

† Ghaffary, S. (March 3, 2026). "Anthropic Nears $20 Billion Revenue Run Rate." Bloomberg.

Metz, C. (2025). The Race to the Top. [historical context on Anthropic's founding and motivations]

2. Legal Analysis and Procurement

† Endrias, M. & Rozenshtein, A. (2026). DPA and supply chain risk analysis. Lawfare.

† OpenAI (February 28, 2026). "Our agreement with the Department of War."

† Axios (March 3, 2026). "Scoop: OpenAI, Pentagon add more surveillance protections to AI deal."

† CNBC (March 3, 2026). "OpenAI's Altman admits defense deal 'looked opportunistic and sloppy.'"

FAR § 9.402(b). Federal Acquisition Regulation. [prohibition on punitive debarment]

Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579 (1952). [constitutional limit on executive power over private industry]

Murthy v. Missouri, 603 U.S. ___ (2024). [government pressure on platforms]

Executive Order 12333 (1981). [the clause enabling surveillance via foreign lines, cited by Masnick]

3. Competitive Landscape: AI and Defense

† Axios (February 23, 2026). "Musk's xAI and Pentagon reach deal to use Grok in classified systems."

† Warren, E. (September 2025). Letter to the Secretary of Defense on the xAI contract.

† eWeek (December 10, 2025). "Pentagon Deploys Google's Gemini AI to 3 Million Workers."

† Military.com (December 10, 2025). "Navy Partners With Palantir for $448 Million AI Submarine Deal."

Allen, G.C. (2017). Artificial Intelligence and National Security. Harvard Kennedy School, Belfer Center.

Payne, K. (2021). I, Warbot: The Dawn of Artificially Intelligent Conflict. Hurst & Co.

Quilty Analytics (2022-2023). Analysis of SpaceX Starlink/DoD contract for Ukraine. [contractual model for purchasing decision rights]

4. AI Sector Economics and Finance

† Epoch AI (February 2026). "Anthropic could surpass OpenAI in revenue by mid-2026."

† Sacra (2026). "Anthropic revenue, valuation & funding." sacra.com.

† Perera, S.A. (February 2026). "The Growth Miracle and the Six Fractures: Anthropic at $380 Billion." Substack.

Toner, H. & Fuchs, G. (2023). AI Investment Dynamics and the Race to Scale. Georgetown CSET.

5. AI Geopolitics: China, Taiwan, TSMC, Semiconductors

† Allen, G. & Thompson, B. (March 5, 2026). Section "Chip Controls and China" in Stratechery interview. [Xi's strategy on TSMC dependency; ZTE/Huawei as one-way door]

Miller, C. (2022). Chip War: The Fight for the World's Most Critical Technology. Scribner.

Beckley, M. & Brands, H. (2022). Danger Zone: The Coming Conflict with China. W.W. Norton.

† Atlantic Council (January 15, 2026). "Eight ways AI will shape geopolitics in 2026."

Khan, S. (2023). Chokepoints: China's Self-Identified Strategic Technology Import Dependencies. Georgetown CSET.

† ESET (January 2026). "Risks in 2026: Espionage, AI hacks, cyber warfare."

6. Disinformation, Military AI, Information Operations

† European Parliament — EPRS (2025). "Information manipulation in the age of generative AI." Briefing.

† EU DisinfoLab (December 2025). "Disinfo Update." disinfo.eu.

† Kuznicka-Blaskowska, D. & Kostyuk, N. (2025). "Emerging need to regulate deepfakes in international law." Journal of Cybersecurity, Oxford.

Bradshaw, S. & Howard, P.N. (2019). The Global Disinformation Order. Oxford Internet Institute.

Marconi, F. (2020). Newsmakers: Artificial Intelligence and the Future of Journalism. Columbia UP.

† Crisis24 (2025). Global Risk Forecast 2026.

7. Scenario Planning Methodology

† Wack, P. (1985). "Scenarios: Uncharted Waters Ahead." Harvard Business Review, No. 85516.

Wack, P. (1985). "Scenarios: Shooting the Rapids." Harvard Business Review, No. 85605.

† Cornelius, P., Van de Putte, A. & Romani, M. (2005). "Three Decades of Scenario Planning in Shell." California Management Review, Vol. 48, No. 1.

† Shell (2025). The 2025 Energy Security Scenarios: Energy and artificial intelligence. shell.com.

Schwartz, P. (1991). The Art of the Long View. Doubleday.

van der Heijden, K. (1996). Scenarios: The Art of Strategic Conversation. Wiley.

8. International Political Economy: Power, Wealth, AI

† Gilpin, R. (1981). War and Change in World Politics. Cambridge University Press.

Gilpin, R. (1987). The Political Economy of International Relations. Princeton UP.

Farrell, H. & Newman, A. (2019). "Weaponized Interdependence." International Security, 44(1).

Nye, J.S. (2021). "Artificial Intelligence and the Problem of Power." PS: Political Science & Politics.

Acemoglu, D. & Johnson, S. (2023). Power and Progress. PublicAffairs.

9. AI Governance, Safety, Constitutional AI

Bostrom, N. (2014). Superintelligence. Oxford UP.

Russell, S. (2019). Human Compatible. Viking.

Anthropic (2022). Constitutional AI: Harmlessness from AI Feedback. https://arxiv.org/abs/2212.08073

Gabriel, I. (2020). "Artificial Intelligence, Values, and Alignment." Minds & Machines, 30.

Raji, I.D. et al. (2020). "Closing the AI Accountability Gap." ACM FAccT 2020.

10. Historical-Legal Context: Government and Private Industry in the U.S.

† Sharp, A. (March 6, 2026). "The End of the World As We Know It." Sharp Text. [source for the following historical precedents]

Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579 (1952). [Truman, steel, constitutional limit on executive power without Congressional authorization; landmark case]

Montgomery Ward & Co. v. United States, 326 U.S. 690 (1946). [FDR, office seizure; government lost; dispute closed as moot when Truman returned the property]

Bamford, J. (2008). The Shadow Factory. Doubleday. [AT&T and Bell Labs as implicit NSA contractors for decades]

Greenwald, G. (2014). No Place to Hide. Metropolitan Books. [AT&T cooperation with PRISM/EO12333]

War Production Board (1942). Orders on Ford, GM, DuPont. National Archives. [forced wartime industrial conversion precedents]

Hounshell, D. & Smith, J.K. (1988). Science and Corporate Strategy: DuPont R&D 1902-1980. Cambridge UP. [DuPont monopoly on Allied munitions and subsequent antitrust]