



# OpenAI・国防総省 AI モデル提供合意と Anthropic 排除：米国 AI 軍事利用を巡る激震

## エグゼクティブサマリー

2026年2月、米国防総省（DoD）とAI企業Anthropicの間で、軍事利用におけるAIの安全制限を巡る対立が決定的な局面を迎えた。Anthropicが倫理的ガードレールの撤廃を拒否したことを受け、トランプ大統領は全連邦機関にAnthropic製品の使用停止を命令。国防長官ピート・ヘグセスはAnthropicを「サプライチェーンリスク」に指定した。その直後、OpenAIが国防総省との機密ネットワーク向けAI提供契約を発表し、Anthropicと同様の安全原則を維持しつつ合意に至ったとされる。この一連の出来事は、AI企業の自主規制と国家安全保障の要求が正面衝突した歴史的事例であり、AI産業全体に深い影響を与えている。  
[1][2][3][4]

## 対立の発端：国防総省の「あらゆる合法的」要求

### 契約背景と軍事AI利用の実態

2025年7月、国防総省はAnthropic、OpenAI、xAI、Googleの4社と各約2億ドルのAI提供契約を締結した。Anthropicはこの中で唯一、軍の機密システムでの利用が承認されたAIモデルの提供者であった。Claudeは諜報分析、兵器開発、戦闘作戦といった最高機密レベルの業務に使用されており、Palantir Technologiesを通じて国防総省のシステムに統合されていた。  
[5][6][7][8]

実際に、2026年1月のベネズエラ・マドゥロ前大統領拘束作戦では、米軍の特殊作戦部隊がClaude AIを使用していたことがウォール・ストリート・ジャーナルにより報じられた。カラカスの複数の拠点への空爆を含むこの作戦で83名が死亡したとされるが、Anthropicの利用規約は暴力の促進、兵器開発、監視活動へのClaude使用を明確に禁止していた。  
[9][10]

# 国防総省の要求と Anthropic の拒否

国防総省は 4 社すべてに対し、AI モデルを「あらゆる合法目的 (all lawful purposes)」に使用可能とする条項への同意を求めた。これは兵器開発、諜報収集、戦闘作戦を含む幅広い軍事用途を網羅するものであった。[11][12]

Anthropic が堅持した 2 つの「レッドライン」は以下の通り：[13][12]

- 米国民に対する大規模国内監視への使用禁止
- 完全自律型致死兵器システムへの使用禁止 (適切な人間の判断と管理を欠くもの)

国防総省側は、これらの制限が作戦上非現実的であり、軍事作戦は「グリーゾーン」で展開されるため、民間企業が設定したカテゴリカルな制約は実行不可能であると主張した。ヘグセス国防長官は Anthropic の安全条件を「woke (進歩的すぎる)」と評した。[14][15][13]

## 2026 年 2 月の危機：エスカレーションの全経過

### 交渉決裂からの最後通牒

日付	出来事
2 月 14 日	国防総省が Anthropic 排除を検討中との報道[11]
2 月 16 日	「サプライチェーンリスク」指定の検討が明らかに[16]
2 月 23 日	xAI が Grok の機密システム利用契約を締結[17]
2 月 24 日	ヘグセスがアモデイ CEO と面会、2 月 27 日午後 5 時 1 分を最終期限に設定 [18]
2 月 26 日	アモデイ CEO が「良心に基づき拒否する」と公式声明[5][19]
2 月 27 日	トランプ大統領が全連邦機関に Anthropic 使用停止を命令[2][20]
2 月 27 日	ヘグセスが Anthropic を「サプライチェーンリスク」に指定[4][1]
2 月 27 日	Anthropic が法的に争う意向を表明[4][21]

2月27日	OpenAI が国防総省との機密ネットワーク契約を発表[1][3]
-------	-----------------------------------

## Anthropic の倫理的姿勢

Anthropic CEO ダリオ・アモデイは2月26日、「We cannot in good conscience cede to this request（良心に基づきこの要求に応じることはできない）」と明確に宣言した。アモデイは「戦争の遂行は民間企業ではなく政府が決定する」としつつも、大規模国内監視と完全自律型兵器は「現在の技術の安全で信頼性のある能力の範囲外」とであると強調した。[22]

Anthropic は最大2億ドル規模の政府契約を失うリスクを承知の上で自社方針を貫いた。同社の Responsible Scaling Policy（責任あるスケーリング方針）は、AI システムからの壊滅的リスクを軽減するための安全基準であり、創業以来の中核的方針であった。[13][5]

### 「サプライチェーンリスク」指定の異例さと影響

ヘグセス国防長官が Anthropic に適用した「サプライチェーンリスク」指定は、通常ロシアや中国といった外国の敵対国に関連する企業に対して使用される措置であり、米国の主要テクノロジー企業に適用されるのは前例がない。[4][23]

この指定の影響は甚大である：

- 米軍と取引するすべての請負業者、サプライヤー、パートナーが Anthropic との商取引を行うことが禁止される[21]
- Anthropic の企業顧客の多くは国防総省と何らかの関係を持つ可能性があり、既存の顧客基盤が侵食される恐れがある[24]
- Anthropic の収益の約80%はエンタープライズ顧客によるものであり、これらの顧客が政府との関係を理由に Anthropic 製品から離れるリスクが存在する[25]

Anthropic はこの指定を「legally unsound（法的に不健全）」であり「a dangerous precedent（危険な前例）」であると批判し、法廷で争う意向を表明した。同社は、ヘグセスには Anthropic と取引する企業の軍以外の商取引を制限する法的権限はないと主張している。[4][21]

## OpenAI の戦略的合意：同じ原則、異なるアプローチ

### 合意の構造と内容

OpenAI CEO サム・アルトマンは2月27日夜、国防総省との機密ネットワーク向け AI 展開契約の締結を発表した。アルトマンは X への投稿で以下を明らかにした：[3][26]

「今夜、国防省との合意に達し、機密ネットワークにモデルを展開する。すべてのやり取りにおいて、国防省は安全性への深い敬意と最良の結果を達成するためのパートナーシップへの意欲を示した」[3]

合意の核心的な要素は以下の通り：[27][1][3]

- 「あらゆる合法目的」への使用を許可（Anthropic が拒否した条項）
- 大規模国内監視の禁止と自律型兵器における人間の責任を安全原則として維持
- OpenAI が独自の\*\*「セーフティスタック」\*\*を構築（技術的、政策的、人的制御のシステム）
- モデルがタスクを拒否した場合、政府は準拠を強制しない[27]
- 利用をクラウドシステムに限定し、ドローンや航空機などのエッジシステムでの展開は除外[27]
- OpenAI 社員が政府職員と機密プロジェクトで協力することを許可[28]

## Anthropic と OpenAI のアプローチの違い

項目	Anthropic	OpenAI
「あらゆる合法目的」条項	拒否[12]	受入れ[28]
大規模国内監視禁止	契約上の保証を要求[13]	技術的セーフガードで対応[1]
自律型兵器制限	契約上の保証を要求[13]	技術的セーフガードで対応[1]
制限の実装方法	契約条項（法的拘束力）[13]	セーフティスタック（技術的制御）[27]
エッジデバイス展開	制限を要求	クラウドに限定[27]
交渉開始時期	2025年 から数ヶ月[12]	2月25日（水）から[29]
結果	契約破棄・排除[2]	合意締結[3]

決定的な違いは、Anthropic が**契約条項としての法的保証**を求めたのに対し、OpenAI は**技術的セーフガード**というアプローチを取ったことにある。OpenAI は「合法目的」条項を受け入れつつ、モデル側で特定用途を技術的にブロックする仕組みを提案し、国防総省に受け入れられた。実質的には同じ「レッドライン」を共有しながらも、フレーミングと実装方法が異なった。[15][28]

## アルトマンの政治的・ビジネス的巧みさ

アルトマンは公には Anthropic の安全姿勢を支持し、「Anthropic の設定したレッドラインを共有する」「彼らは本当に安全性を重視している」と CNBC で述べた。社内メモでも「Anthropic と同じレッドラインを持っている」と従業員に伝えていた。[30][22]

同時に、裏では 2 月 25 日水曜日から国防総省と交渉を開始し、Anthropic が排除された当日に契約締結を発表するという、ビジネスと政治の両面で巧みな立ち回りを見せた。NY タイムズはこれを「ビジネスと政治の両面での勝利」と評した。[29]

## 競合他社の動向：xAI と Google の立場

### xAI (Grok)

イーロン・マスクの xAI は 2 月 23 日、Grok モデルの機密システム利用契約を国防総省と締結した。xAI は「あらゆる合法目的」条項に同意し、Anthropic に次いで 2 社目の機密環境での利用承認を獲得した。ただし、国防総省関係者は Grok が Claude ほどの性能はなく、「より正確な情報を一貫して提供する」Claude の代替は容易ではないと認めている。[17][31][32][5]

### Google (Gemini)

Google は 2025 年 12 月に GenAI.mil プラットフォームとして Gemini を国防総省の約 300 万人の文民・軍人向けに提供する契約を獲得していた。2 月時点で機密システムへの展開についても合意に近いとされているが、最終契約には至っていなかった。[6][32]

## Anthropic への事業影響と IPO 計画

### 財務的インパクト

Anthropic は 2026 年初頭時点で企業価値約 3,500 億ドル、年間売上ランレートが 90 億ドルを超え、2026 年には最大 260 億ドルに達する見込みであった。200 億ドル超の新規資金調達ラウンドも進行中であった。[33][25]

2 億ドルの国防総省契約そのものは売上全体に対して小さいが、「サプライチェーンリスク」指定の波及効果は桁違いに大きい。軍と取引するあらゆる企業が Anthropic との商取引を証明・排除する必要が生じるため、エンタープライズ顧客基盤（売上の 80%）が浸食されるリスクがある。[24][25]

### IPO 計画への影響

Anthropic は 2026 年中の IPO を視野に、法律事務所 Wilson Sonsini を起用し準備を進めていた。評価額は 3,000 億ドル超となれば、史上最大級の AI 関連 IPO となるはずであった。しかし、連邦政府からの排除と「サプライチェーンリスク」指定は、投資家心理と企業顧客の離反リスクに直結し、IPO の時期や評価額に大きな不確実性をもたらす可能性がある。[34][35]

# 法的・倫理的論点

## 国際人道法と AI 安全ガードレール

Anthropic が求めたガードレールは、しばしば倫理的コミットメントとして描写されるが、その実質は法的リスク管理に直結している。同社が契約に盛り込もうとした 3 つの条件は：[13]

1. 米国市民に対する大規模監視への不使用
2. 適切な人間の判断と管理を欠く自律型兵器システムの開発・展開への不使用
3. 明確な人間の監督なしの攻撃的・致死的意思決定への不使用

これらは既存の法的枠組み（国際人道法における区別原則・均衡原則）に直接対応するものであり、Anthropic 独自の「過激な」要求ではないという見方がある。実際に、国防総省自身も AI をこれらの目的に使用する意図はないと述べていたが、民間企業がその条件を設定することには反対していた。[36][18][13]

## 国防生産法の援用可能性

国防総省は、朝鮮戦争時代の\*\*国防生産法（Defense Production Act）\*\*を援用して Anthropic に強制的に協力させることも検討していた。この法律は大統領に国内産業を監督する広範な緊急権限を付与するものであり、AI 企業に適用されれば前例のない展開となる。[30]

## 前例としての危険性

フォーブスの分析によれば、国防総省が Anthropic を「サプライチェーンリスク」に指定した行為は、すべての米国企業に対する警告となる。従来この指定は外国の敵対的企業に対して使用されてきたものであり、米国の主要テクノロジー企業に適用されたのは初めてである。これは「政府と交渉する際に自社のポリシーを主張すれば、敵国扱いされるリスクがある」という前例を作りかねない。[23][4]

# AI 軍事利用の今後と産業への示唆

## テクノロジー業界への波及

この紛争は、フロンティア AI 企業が国家の軍事要求にどう対応するかという根本的な問題を提起している。注目すべきは、OpenAI のアルトマンが Anthropic の安全姿勢を公に支持しながらも、異なるアプローチで合意に至った点である。これは「原則を共有しつつ実装方法で差別化する」というモデルが、今後の AI 企業と政府の関係における新たなテンプレートとなる可能性を示唆している。[15][30]

欧州への波及も懸念されている。英国を含む複数の欧州諸国がこれらの AI 企業と大規模契約を結んでおり、米国での先例が国際的な安全基準の議論に影響を与える可能性がある。[13]

## IP・テクノロジー戦略への含意

この事例は、AI企業の利用規約（Acceptable Use Policy）と安全方針が、単なるコンプライアンス文書ではなく**事業存続に関わる戦略的資産**であることを鮮明にした。Anthropicのケースは以下の教訓を示している：

- **安全方針の設計**は、倫理と事業戦略の交差点における最も重要な意思決定の一つとなった
- **契約条項 vs 技術的制御**という実装の選択が、同じ原則に基づいても全く異なる結果をもたらさうる（Anthropic vs OpenAI）
- AIモデルの**軍事利用に関する特許・知的財産**の保護と活用は、政府との関係によって大きく左右される

## 今後の焦点

Anthropicが「サプライチェーンリスク」指定に対して提起する法廷闘争の行方が、今後のAI産業と政府の関係を規定する重要な判例となる。同時に、OpenAIの「セーフティスタック」アプローチが実際にどの程度機能するか、技術的にどのように大規模監視や自律型兵器への転用を防ぐのかについての検証も求められる。この問題は、AI技術のガバナンスにおける最も困難かつ重要な課題として、今後も議論が続くことになるだろう。[37][21]

---

## References

1. [OpenAI strikes deal with Pentagon hours after Trump admin bans Anthropic](#) - OpenAI CEO Sam Altman announced late Friday that the company had signed a deal with the Pentagon for...
2. [Trump orders federal agencies to 'immediately cease' using Anthropic technology](#) - President Trump on Friday directed federal agencies to 'immediately cease' using Anthropic technolog...
3. [OpenAI announces new deal with Pentagon — including ethical ...](#) - “Tonight, we reached an agreement with the Department of War to deploy our models in their classifie...
4. [Anthropic to challenge Pentagon's supply chain risk tag](#) - The Hill - Anthropic, an AI company, disputes the Pentagon's decision to label it a supply chain risk, calling ...
5. [Anthropic says it 'cannot in good conscience' allow Pentagon to remove AI checks](#) - Pete Hegseth had threatened to cancel \$200m contract unless it was given unfettered access to Claude...
6. [Pentagon picks Google AI platform for its millions of employees](#) - 'The future of American warfare is here, and it's spelled AI,' Defense Secretary Pete Hegseth said i...

7. [US military used Anthropic's Claude AI in Nicolás Maduro capture ...](#) - The U.S. military reportedly used Anthropic's AI tool Claude in the operation that captured Venezuel...
8. [US used Anthropic's Claude during the Venezuela raid, WSJ reports](#) - Anthropic's artificial-intelligence model Claude was used in the U.S. military's operation to captur...
9. [US military used Anthropic's AI model Claude in Venezuela raid, report says](#) - Wall Street Journal says Claude used in operation via Anthropic's partnership with Palantir Technolo...
10. [Pentagon Used Anthropic's Claude in Maduro Venezuela Raid - WSJ](#) - Anthropic's artificial-intelligence tool Claude was used in the U.S. military's operation to capture...
11. [Pentagon threatens to cut off Anthropic in AI safeguards dispute, Axios reports](#) - The Pentagon is considering ending its relationship with artificial intelligence company Anthropic o...
12. [Pentagon threatens to cut off Anthropic in AI safeguards dispute](#) - Anthropic has not agreed to the Pentagon's terms and defense officials are getting fed up after mont...
13. [The Pentagon/Anthropic Clash Over Military AI Guardrails](#) - Just this week, a contract dispute between the U.S. Department of Defense (DoD) and the AI company A...
14. [The Pentagon/Anthropic Clash Over Military AI Guardrails](#) - [Jessica Dorsey is an Assistant Professor of International Law at Utrecht University School of Law; ...
15. [Pentagon approves OpenAI safety red lines after dumping Anthropic](#) - The Pentagon has complained Anthropic's red lines on military use were "woke."
16. [The Department of Defense is considering designating Anthropic as ...](#) - The Pentagon's relationship with Anthropic has deteriorated over the military use of AI , and it has...
17. [Musk's xAI, Pentagon reach deal to use Grok in classified systems](#) - Elon Musk's artificial intelligence company xAI has signed an agreement to allow the military to use...
18. [Pete Hegseth meets with Anthropic CEO over disagreements about AI guardrails for military use](#) - Defense Secretary Pete Hegseth gave Anthropic CEO Dario Amodei a Friday deadline to comply with dema...
19. [Anthropic rejects latest Pentagon offer: 'We cannot in good ...](#) - CNN - Anthropic is rejecting the Pentagon's latest offer to change their contract, saying the changes do n...
20. [Trump Moves to Ban Anthropic From the US Government](#) - President Donald Trump's sudden order comes after the Defense Department pressured Anthropic to drop...

21. [Anthropic says will challenge 'supply chain risk' label in court after US bars agencies from using its product](#) - The AI firm began 2026 on a winning streak, with its sales surging, multiple products going viral an...
22. [OpenAI announces Pentagon deal after Trump bans Anthropic](#) - NPR - President Trump ordered the U.S. government to stop using the artificial intelligence company Anthro...
23. [With Us Or Against Us: The Pentagon's Warning To American Businesses](#) - The Pentagon is threatening to designate Anthropic a "supply chain risk" for keeping AI safety guard...
24. [The clock is ticking down on a critical Pentagon deadline for Anthropic](#) - American AI company Anthropic has until 5:01 pm ET to give in to the Pentagon's demands or face bein...
25. [Anthropic IPO 2026: Latest Timeline, Valuation, and Risks](#) - Anthropic is reportedly preparing for a possible 2026 IPO. Here is the latest Anthropic IPO timeline...
26. [OpenAI strikes deal with Pentagon, hours after rival Anthropic was blacklisted by Trump](#) - OpenAI CEO Sam Altman said on X that his company has come to terms with the Defense Department for u...
27. [Agreement reached with Department of War to deploy OpenAI models in classified network: CEO Sam Altman](#) - Sam Altman said OpenAI has agreed with the Department of War to run its AI models on the classified ...
28. [OpenAI Reaches A.I. Agreement With Defense Dept. After Anthropic ...](#) - These moves allowed Mr. Altman to uphold safety principles around A.I. while still landing the Penta...
29. [OpenAI reaches AI agreement with Pentagon after Anthropic clash](#) - The deal came hours after President Trump had ordered federal agencies to stop using artificial inte...
30. [OpenAI says it shares Anthropic's 'red lines' over military AI use](#) - OPB - OpenAI's Sam Altman says he shares the "red lines" set by rival Anthropic restricting how the milita...
31. [Pentagon, Musk's xAI reach agreement to use Grok in classified ...](#) - New agreement allows xAI to deploy its Grok AI model within classified military systems amid tension...
32. [Pentagon Gives A.I. Company an Ultimatum](#) - NYTimes.com - The Pentagon also is close to an agreement with Google to bring its Gemini model onto the classified...
33. [Anthropic's More Than \\$20 Billion Funding to Close as Soon as Next ...](#) - The latest funding round would nearly double Anthropic's prior valuation and comes just five months ...
34. [Anthropic quietly makes moves toward blockbuster IPO](#) - LinkedIn - Anthropic's IPO could be one of the biggest ever, valuing the company at more than \$300 billion — an...

35. [Anthropic plans an IPO as early as 2026, FT reports - Reuters](#) - Dario Amodei-led Anthropic is projecting to more than double and potentially nearly triple its annua...
36. [Trump bans AI firm Anthropic from federal agencies](#) - The White House had demanded Anthropic to allow unrestricted military use of its AI technology or fa...
37. [Anthropic says it will challenge Pentagon's supply chain risk designation in court](#) - Feb 27 (Reuters) - Anthropic said on Friday it will challenge in court the Pentagon's decision to de...