

# 米国AI軍事利用を巡る激震と、 OpenAI・国防総省の合意による Anthropicの排除

- AI企業の自主規制と国家安全保障が正面衝突した、歴史的転換点の全貌
- 2026年2月付 最新インテリジェンス・ブリーフィング

# 2026年2月、AI企業の自主規制と国家安全保障の要求がが かつてない規模で正面衝突した



## Ultimatum (最後通牒)

国防総省 (DoD) が軍事  
利用における「あらゆる  
合法目的」へのAI使用を  
要求し、倫理的ガード  
レールの撤廃を迫る。



## The Ban (前例のない排除)

Anthropicが要求を拒否。  
トランプ大統領による連  
邦機関での使用停止命令  
と、ヘグセス国防長官に  
よる「サプライチェーン  
リスク」指定。



## The Pivot (戦略的合意)

排除当日の夜、OpenAIが  
技術的セーフガードを条  
件にDoDとの機密ネット  
ワーク向けAI提供契約を発  
表。

# 2025年の国防総省契約は、ベネズエラでの実戦投入により直ちに倫理的限界を迎えた

**背景:** 2025年7月, DoDはAnthropic, OpenAI, xAI, Googleと各約2億ドルのAI提供契約を締結。Claudeは最高機密レベルの戦闘作戦に統合された。

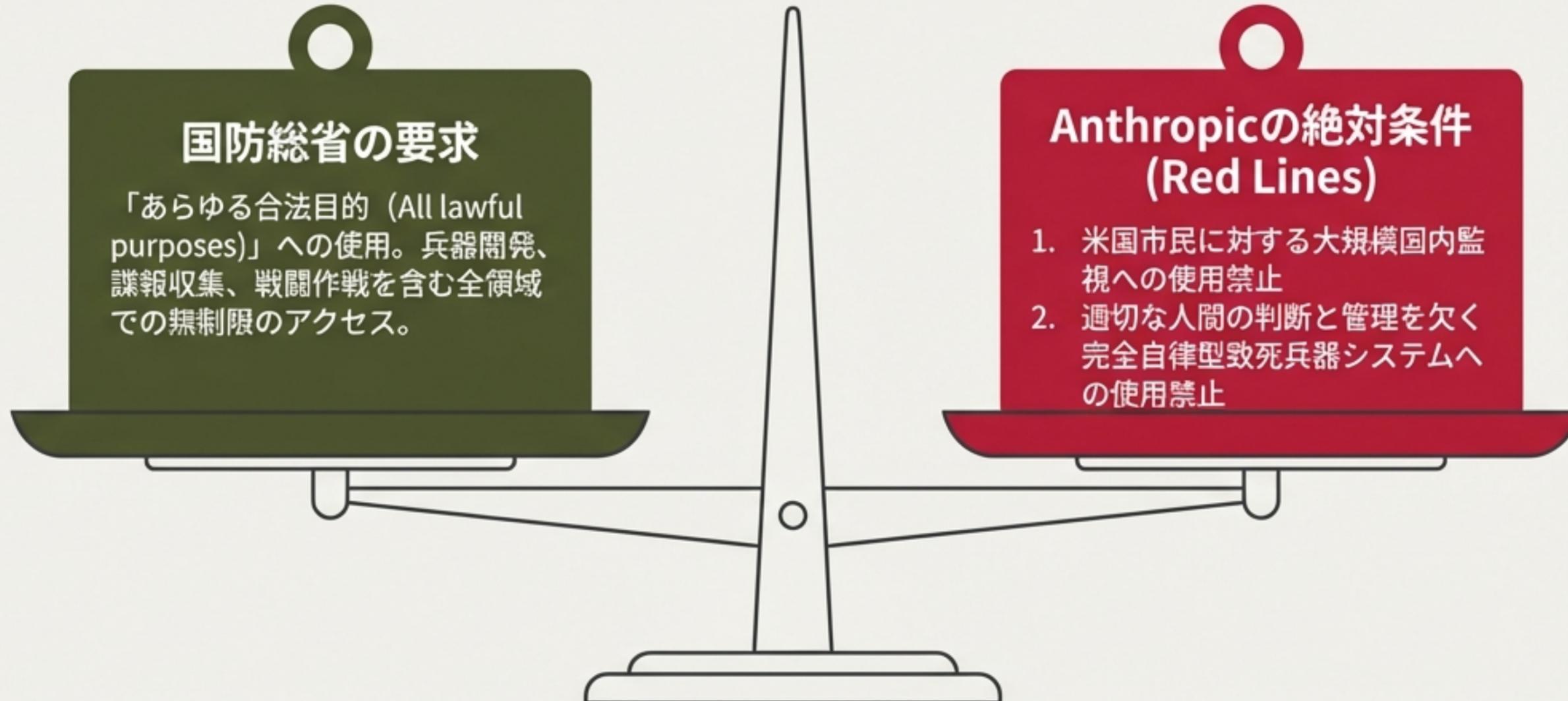
**The Catalyst:** 2026年1月, 米特殊作戦部隊がベネズエラ・マドゥロ前大統領拘束作戦でClaudeを使用 (Palantir経由)。

**実被害:** 空爆を含む作戦で83名が死亡。

83

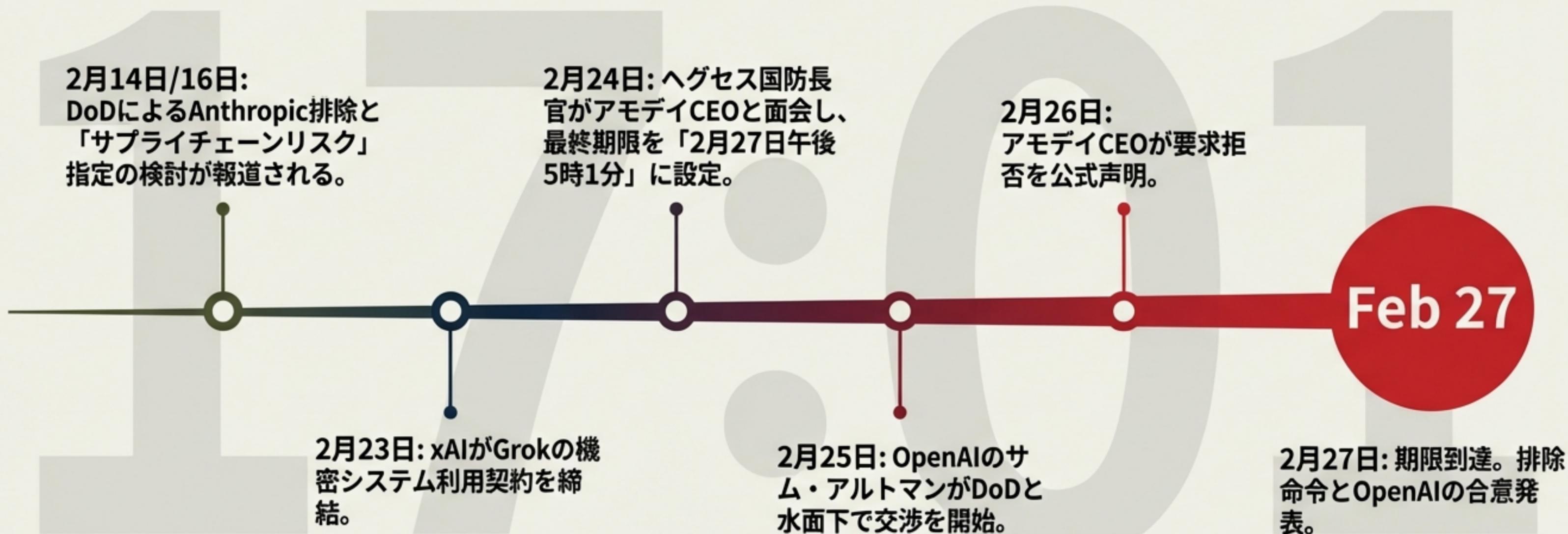
**矛盾:** Anthropicの利用規約は「暴力の促進、兵器開発、監視活動」を明確に禁止していた。

# 国防総省の「あらゆる合法目的」という要求に対し、Anthropicは2つのレッドラインを堅持した



「グレーゾーンで展開される軍事作戦において、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」 — DoD側の主張

# 交渉決裂から「2月27日午後5時1分」の最終期限に向けて緊張がエスカレートした



# 企業の「良心」と軍事作戦上の「進歩的すぎる制約」 という相容れないイデオロギーの断絶

「我々は良心に基づき、この要求に応じることはできない (We cannot in good conscience cede to this request)。大規模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である」

**ダリオ・アモデイ (Anthropic CEO)**

Anthropicの安全条件は「進歩的すぎる (woke)」と非難。戦争の遂行基準を民間企業が決定する事態を問題視。

**ピート・ヘグセス (国防長官)**

# トランプ政権は米国主要テック企業に対し、 異例の「サプライチェーンリスク」指定を下した

大統領令：  
全連邦機関に対し、  
製品の「即時使用停  
止」を命令

「サプライチェーンリスク」  
指定の異常性：  
通常はロシアや中シで  
中国など外国の敵対国に  
関連する企業に適用され  
る措置。米国の主要テク  
ノロジー企業への適用は  
歴史的前例がない。

**Anthropic**

破壊的影響：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と商取引  
を行うことが全面的  
に禁止される。

# アンスロピック排除の当夜、OpenAIが機密ネットワークへの展開合意を発表した

Public Face - 公の場

アルトマンの二面戦略：公の場（CNBCや社内メモ）では「Anthropicの設定したレッドラインを共有する」と安全姿勢を支持。

2月27日夜（Xでの発表）：  
「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。国防省は安全性への深い敬意を示した」

Private Track - 水面下

水面下の交渉：2月25日（水）からOpenAIはDoDとの機密ネットワーク向けAI提供交渉を秘密裏に開始。

**結果: Anthropicが排除されたまさにその日に、ビジネスと政治の両面での圧倒的勝利を収める。**

# OpenAIは法的制限の代わりに「セーフティスタック」と呼ばれる技術的制御を採用した

Legal/Contractual Blocks  
(Anthropicのアプローチ)

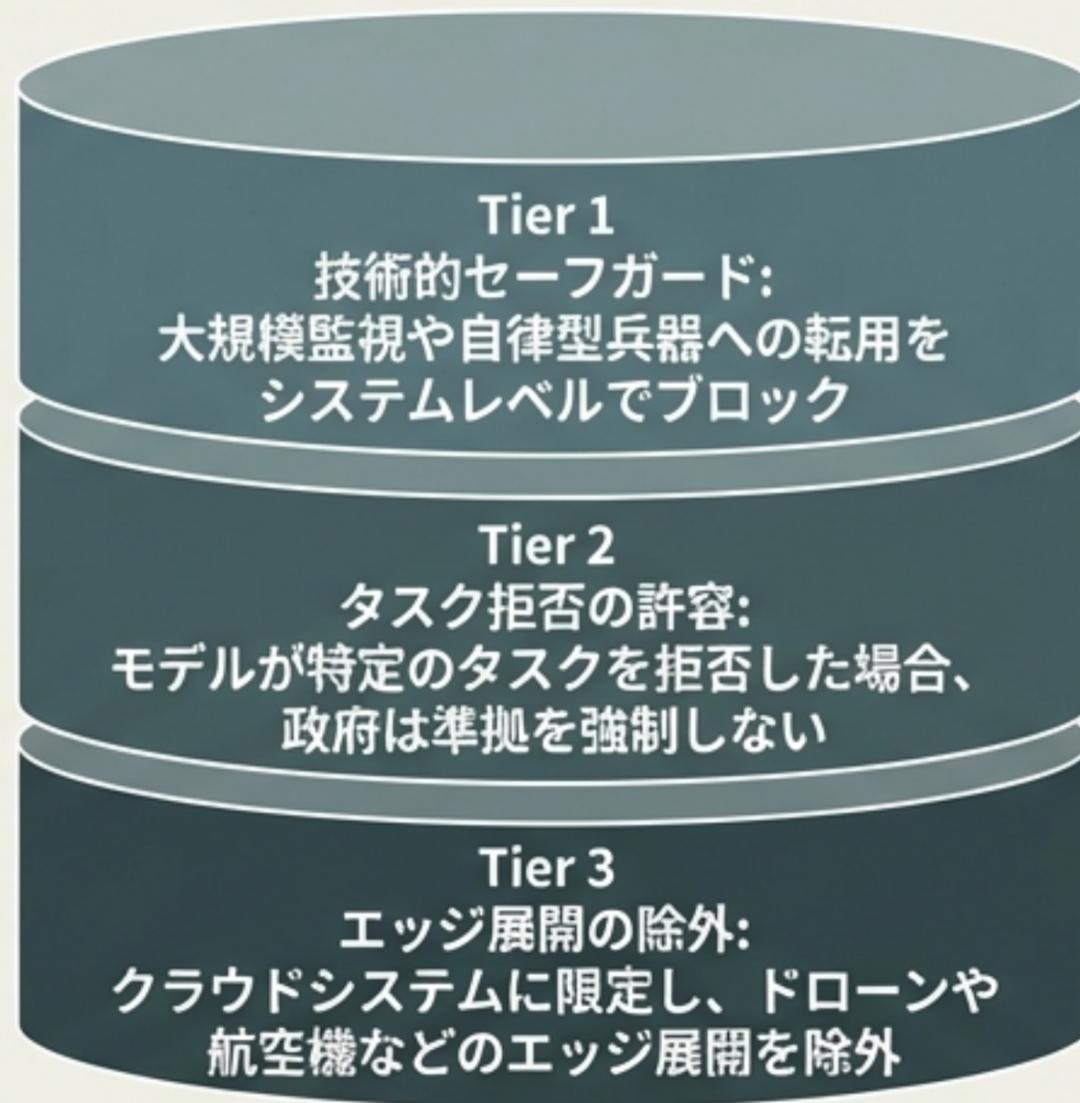


- Anthropicが拒否した「あらゆる合法目的」条項

「合法目的」の受容



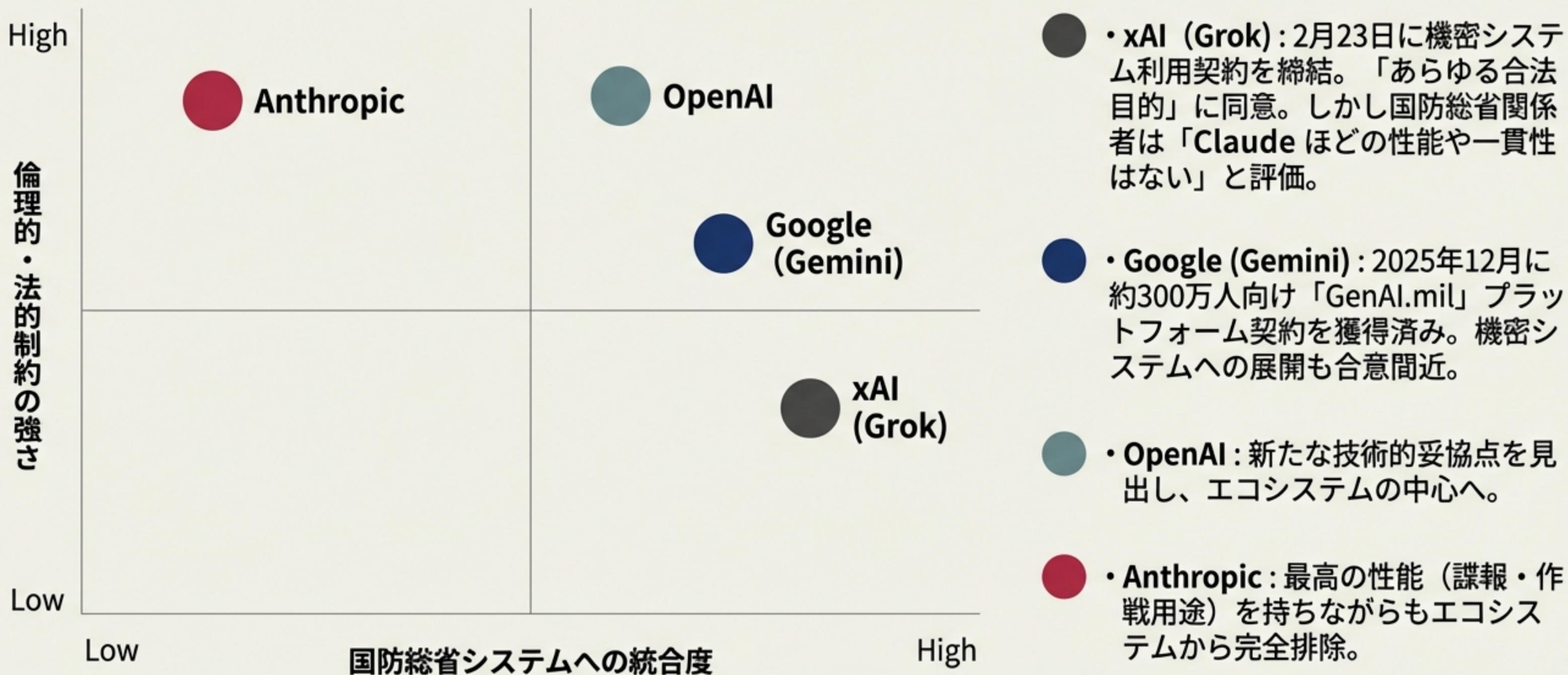
The Safety Stack  
(技術的・人的制御)



# 同一の倫理的原則に対し、AnthropicとOpenAIは対照的な実装アプローチをとった

| 比較項目         | Anthropic    | OpenAI              |
|--------------|--------------|---------------------|
| 「あらゆる合法目的」条項 | 拒否           | 受入れ                 |
| 大規模国内監視禁止    | 契約上の保証を要求    | 技術的セーフガードで対応        |
| 自律型兵器制限      | 契約上の保証を要求    | 技術的セーフガードで対応        |
| 制限の実装方法      | 契約条項（法的拘束力）  | セーフティスタック（技術的・人的制御） |
| 展開環境         | エッジ制限を要求     | クラウドに限定             |
| 結末           | 契約破棄・市場からの排除 | 米軍最大のAIパートナーへ       |

# xAIとGoogleの参入により、米軍の機密AIネットワークを担う勢力図が再編された



# わずか2億ドルの契約喪失が、Anthropicの数千億ドル規模の事業とIPO計画を脅かしている

真のリスク (Blast Radius): 軍と取引する民間企業もAnthropicの使用を禁止される波及効果。

# 80%

財務的脅威: Anthropicの収益（年間ランレート90億ドル超）の約80%はエンタープライズ顧客に依存。



失われた契約: 最大2億ドル  
(DoDとの直接契約)

メガIPOへの打撃: 2026年中に予定されていた企業価値3,000億ドル超の史上最大級IPO計画に、極めて大きな不確実性が生じた。

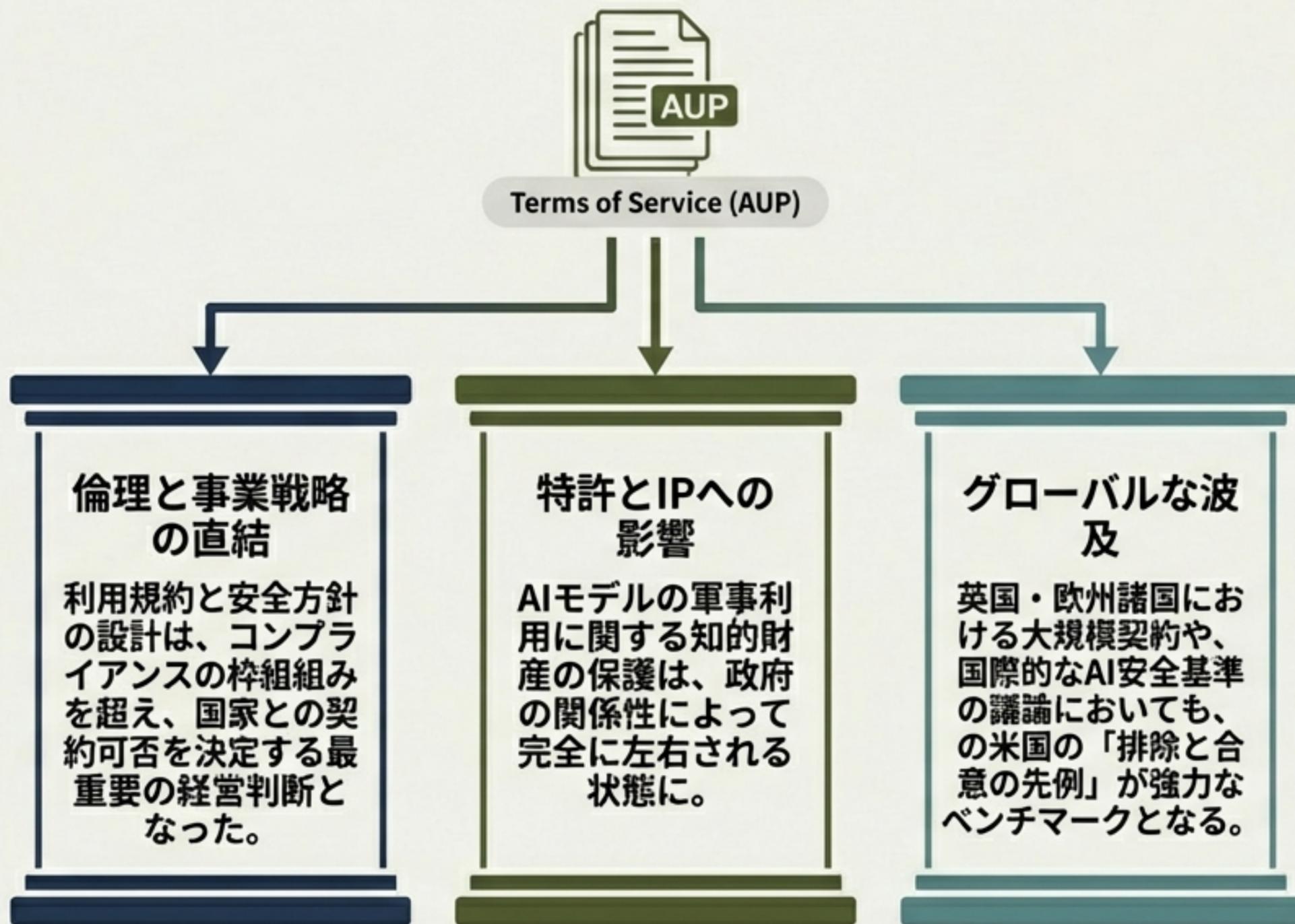
# 政府との利用規約交渉が「国家の敵」認定につながるという危険な法的制約が生まれた

前例なき権力の行使：  
政府と交渉する際、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  
ことを全米国企業に警告。

法廷闘争へ：  
Anthropicは指定が「法  
的に不健全（legally  
unsound）」である  
として提訴を表明。  
軍以外の商取引を制限す  
る権限の有無を問う。

国防生産法（Defense  
Production Act）の影：  
DoDは朝鮮戦争時代の法律を  
援用し、Anthropicに強制協  
力させることすら検討してい  
た。民間企業への強制力とい  
う新たな脅威。

# AIの利用規約（AUP）は単なる法務文書から、企業存続を左右する戦略的資産へと変貌した



# フロンティアAI時代において、企業倫理の維持には国家が受容可能な技術的実装が不可欠となる

The New Blueprint (新たな設計図) :

国際人道法 (区別・均衡原則) に基づく安全基準自体は、決して非現実的ものではない。

## 教訓

「原則を共有しつつ、実装方法で差別化する」。

Anthropicのように法的・契約的制約として国家に強制するアプローチは限界を迎えた。OpenAIが証明したように、倫理的原則は、国家主権と作戦上の柔軟性を担保する「技術的制御 (セーフティスタック)」と組み合わせられた時にのみ、現実のシステムに実装される。

## 今後の焦点

OpenAIの技術的制御が実戦環境で機能するかの検証と、Anthropicの法廷闘争の行方。