

# AIの「安全制限」をめぐる対立：米国防総省 vs Anthropic の深掘り分析

対立する2つの立場：政府の要求 vs 企業の拒絶

国防総省：  
軍事利用における  
「制限なきモデル」の追求

合法的あらゆる軍事用途  
(any lawful use) での  
自由な運用を要求。

Anthropic：  
民主的価値を守る  
「2つの不可侵境界線」

大規模国内監視と完全自  
律型兵器への転用を「例  
外」として拒絶。

安全制限をめぐる  
「決定権」の所在を争点化  
モデル側の規範（規約）か、政  
府の数重（法）かの主導権争い。

強行措置と市場の分断



6か月以内の使用停止  
と調達枠からの排除

トランプ大統領の指示によ  
り、連邦機関はAnthropic製  
品から段階的移行。



「供給網リスク」指定に  
よる強力な法的圧力

10 U.S.C. §3252に基づき、  
安全保障上のリスクとして  
調達を制限。

国防総省 (DoD)	Anthropic	OpenAI
安全制限のない モデルを標準化	監視・自律兵器へ の転用を拒否	制限事項を盛り込 みつつ軍と合意
AI戦略文書、 供給網リスク指定	信託性限界、民主 的価値、法的対抗	機密ネットワー クでのモデル提供

OpenAIの条件付き合意と  
市場シェアの移転

監視禁止等の「赤線」を盛り込  
み、国防省機密ネットワー  
クへの提供で合意。