GLOBAL
EQUITY

2026. 6. 15

글로벌주식팀
이영진 Senior Analyst, CFA
youngjin91.lee@samsung.com

리서치센터 리포트
바로가기

글로벌 AI

앤스로픽의 Claude Fable 5와 Mythos 5 비활성화 결정

- 앤스로픽은 미국 정부의 수출통제 정책 준수를 위해 모든 고객에게 최신 모델 비활성화 결정. 양 측의 의견이 엇갈리고 있지만, 단기에 앤스로픽에 부정적 이슈
- 프론티어 모델이 전략자산으로 취급. 향후 각 국의 소버린 AI 전략 강화 전망

WHAT'S THE STORY?

AI 모델의 상징적 발전 후퇴 사례: 미 정부가 국가안보 권한 근거로 앤스로픽의 Claude Fable 5와 Mythos 5 모델에 대한 모든 외국인의 접근을 중단하는 수출통제 지침 발령. 앤스로픽은 규정 준수를 위해 모든 고객에 대한 모델 접근 즉시 중단. 다만 앤스로픽의 타 모델에 대한 접근은 영향을 받지 않는다고 언급.

향후 진행 상황 확인이 필요하지만, 이번 조치는 특정 모델의 일시적 비활성화가 아니라, 프론티어 AI 모델의 상업적 배포권이 국가안보와 수출통제 영역으로 편입되기 시작했음을 시사. AI 기업의 경쟁력은 단순 모델 성능, 컴퓨트 확보, 제품화 역량뿐 아니라 국가와 고객에게 모델을 안정적으로 제공할 수 있는지에 의해서도 좌우

이번 사안의 핵심은 네 가지. 1) 프론티어 모델이 단순 소프트웨어가 아니라 수출통제 가능한 전략자산으로 취급되기 시작, 2) 모델 성능뿐 아니라 배포 가능 국가, 사용자, 산업을 통제하는 라이선스 권한이 개별 기업의 밸류에이션 변수로 부상, 3) 외국인 직원 접근 제한은 AI 기업 내부 연구조직 설계와 개발 속도에 직접적 부담으로 작용 가능, 4) 각국의 소버린 AI(자국어 모델, 공공 및 국방용 폐쇄 모델 등) 강화 가능성

일련의 타임라인: 추가 보도에 따르면 목요일 아마존을 비롯한 다수 기업이 신규 모델의 안전장치 우회 가능성에 대한 우려를 백악관에 전달. 해당 사안은 금요일 아침 백악관 최고위층까지 보고됐고, 백악관은 NSA 등과 협력해 국가안보 차원의 위험 여부 평가. 베센트 재무장관, 케언크로스 백악관 사이버 담당 책임자, 수지 와일스 비서실장 등이 주요 고위직이 참석한 긴급 회의도 진행

앤스로픽의 다리오 아모데이와 정부 고위 관계자는 총 3차례 통화 진행. 앤스로픽 측은 정부가 문제를 오해하고 있다고 판단. 발견된 사례는 특정 방식의 제한적 우회 사례일 뿐이며, 범용적 탈옥(Universal Jailbreak)이나 안전장치의 전면 무력화는 아니라고 주장. 이에 추가 정보와 대응 시간을 요청한 것으로 알려짐

하지만 백악관 측은 아모데이의 설명에 설득되지 않았고, 발견된 사례를 국가안보 차원의 위험으로 최종 판단. 모델의 자발적 철수 및 공동 대응을 요구했으나, 앤스로픽의 대응이 충분하지 않다고 판단해 수출통제 조치 발동. 이 과정에서 앤스로픽과 정부는 1) 발견된 취약점의 심각성, 2) 앤스로픽 대응과 정부 조치 방식, 3) 출시 전 안전성 검증의 적절성 등을 두고 엇갈리는 주장을 이어가는 상황

(다음 페이지에 계속)

안전성 담론의 역설과 앤스로픽에 대한 불신 분위기: 최근 아모데이는 "Policy on the AI Exponential"라는 블로그 글을 통해 프론티어 AI 모델의 위험성을 강조. 제3자의 위험 의무평가, 정부의 모델 배포 차단 및 철회 권한, 보고 의무 강화 등을 주장. 결과적으로 안전성 규제를 주장해온 앤스로픽이 규제 논리의 첫 적용 대상이 된 아이러니 발생. 앤스로픽이 강조해온 안전성 담론이 역설적으로 정부 개입의 정당화 논리로 활용된 셈

또한 행정부는 기존에도 앤스로픽이 최신 모델의 보안 위험을 적절히 관리할 수 있을지에 대한 의문을 보유. 특히 아모데이와 통화에서 앤스로픽이 정부 보안 전문가와 협력해 문제를 해결할 의사가 충분하지 않다는 신호로 해석. 정부는 앞선 국방부 갈등과 이번 조치가 별개의 모델 안정성 이슈라는 입장. 다만 앤스로픽의 바이든 행정부 출신 인사 다수 채용, 아모데이의 트럼프 행정부 비판 등은 행정부 내부에서 앤스로픽에 대한 불편한 시선의 일부 요인으로 작용했을 가능성

아마존은 왜?: 앤스로픽의 주요 투자자이자 파트너인 아마존이 이번 조치의 도화선이 되었다는 보도 내용. 아마존 연구진은 특정 프롬프트 조합을 통해 Fable 5가 차단해야 할 사이버 공격 관련 정보를 제공하도록 유도하는 데 성공. 최소 4개의 소프트웨어 프로그램에서 보안 취약점 탐색에 성공했고, 앤디 재시 CEO가 이를 정부 관계자에게 설명한 상황으로 이해

해당 정보는 Fable 5의 안전장치에 의해 제공되면 안 되는 정보지만, 실제 사이버 공격 정보라고 보기에는 애매하다는 의견도 존재. 아마존 연구진이 발견한 취약점을 익스플로잇 코드로 변환하는 단계까지 접근했다는 증거도 부재하다는 분석. 아마존 대변인은 대규모 민간 및 공공 부문 고객을 지원하는 주요 클라우드 기업으로서 정부가 잠재적 보안 리스크에 대한 자문을 구하는 경우는 드문 일이 아니라는 입장

AI 산업에 대한 정부의 통제력 강화: 이번 수출통제 조치뿐 아니라 최근 트럼프 대통령은 AI 모델 감독 권한 확대 관련 행정명령에도 최종 서명. 정부의 AI 기업 지분 보유 방안 논의 관련 보도도 이어지는 상황. 다만 백악관 관계자는 이번 앤스로픽 수출통제 조치가 앤스로픽의 대응 부족에 따른 개별 사안이며, 다른 AI 기업으로 확대될 가능성은 낮다는 입장

앤스로픽의 매출과 조직 내부에 부정적 이슈: 앤스로픽 입장에서는 단기적으로 매출 관련 부정적 영향 존재. 막대한 비용을 들여 개발한 모델이 미국 외 지역에서 매출을 발생시키는 데 제약 발생 또한 내부 외국인 연구진의 접근 제한도 부담. 프론티어 AI 랩 내 외국인 연구진 비중은 대체로 두 자릿수 수준으로 알려져 있으며, 최근 합류한 핵심 연구진 중 일부도 미국 국적이 아닌 점이 부각

이에 따라 프론티어 모델 개발 과정에서 국적 기반 접근통제, 연구 권한 분리, 컴플라이언스 체계 구축 필요성이 커질 가능성. AI 기업 내부 개발 구조 내 별도 체계 필요로 연결. 결국 연구 속도 저하, 조직 설계 비용 증가, 컴플라이언스 부담 확대, 내부 사일로 비대화로 귀결 가능

정부 관계가 새로운 리스크 요인이자 밸류에이션 변수로 부상. 앤스로픽은 모델 성능을 기반으로 국방부와의 갈등을 돌파하는 듯했으나, 오히려 높은 성능이 규제 요인으로 작용. 엔터프라이즈 고객 입장에서도 모델 성능뿐 아니라 공급 안정성과 정책 리스크를 함께 평가할 필요.

오픈AI 입장에서 반사이익 기대 가능. Mythos 5와 Fable 5가 막혀 있는 사이 GPT-5.5로 기존 경쟁하던 엔터프라이즈 시장 공략 가능. 국방부와 앤스로픽 간 갈등 속에서도 오픈AI는 친 정부적 전략을 표방. 물론 앤스로픽이 공식 성명에서 GPT-5.5에서도 유사한 취약점이 발생할 수 있다는 주장을 펼친 만큼 향후 정부 통제의 확산 추이는 확인 필요

당연하게도 소버린 AI 전략 부상: 소버린 AI는 정치적 구호처럼 느껴졌지만, 이제는 현실적 필요성으로 부상. 미, 중이 프론티어 모델 경쟁을 이어가는 가운데, 그 외 지역은 선택의 갈림길에 놓여 있음. 동맹국의 프론티어 모델을 활용하는 전략이 언제든 정책적으로 막힐 수 있다는 가능성이 대두

각국은 해외 모델에 대한 종속을 피하는 전략을 강화할 것. 특히 자국 내 공공 및 안보 시스템에 활용되는 모델이 수출통제 같은 정책 변수에 좌우되지 않는 것이 중요해질 전망. 핵심 산업용 AI에서도 국내 통제가 가능한 대체재 탐색이 확대. 중장기적으로 오픈웨이트, 자국어 특화 모델, 정부 및 국방용 폐쇄 모델 수요 증가 전망. 또한 모델 주권뿐 아니라 데이터, 클라우드, 컴퓨트 주권의 중요성도 함께 강조될 것. 다만 자체 프론티어 모델 개발 전략은 탑티어 성능 경쟁 여부와 자본 측면에서 여전히 난이도가 높은 선택

중국 및 오픈웨이트 진영의 기회와 한계: 미니맥스와 Z.ai 등 중국 모델 기업들은 사태 이후 자신들의 모델 정책이 백악관 결정과 반대된다는 점을 강조. 미국 프론티어 모델의 접근 안정성이 흔들릴 경우, 중국 및 오픈웨이트 모델 진영은 배포 지속성과 정책 독립성을 차별화 포인트로 활용 가능. 다만 중국 모델은 미국 기술 증류 의혹, 규제 리스크, 신뢰성 문제를 동시에 안고 있어 공공 및 안보 영역의 대체재가 되기에는 한계 존재. 따라서 단기적으로는 미국 모델의 진정한 대체재라기보다, 각국의 멀티모델 전략과 자체 통제 가능한 AI 스택을 강화하는 계기로 보는 것이 적절

구조적인 함의: 이번 사안은 앤스로픽의 단기적인 부정적 영향보다 프론티어 모델이 더 이상 단순 API 제품이 아니라 국가가 통제할 수 있는 전략자산으로 인식되었다는 구조적 함의 내포. 프론티어 AI 기업의 경쟁력도 이제 모델을 얼마나 잘 만들 수 있는지가 아니라, 모델을 어디에, 누구에게, 얼마나 안정적으로 배포할 수 있는지까지 포함하는 문제로 확장

AI 기업의 밸류에이션에도 정부와의 관계, 외국인 연구진 접근 관리, 국가별 라이선스 구조, 컴플라이언스 비용 등이 반영. 동시에 미국 모델 의존도가 높은 국가와 기업 입장에서는 소버린 AI, 자국어 특화 모델, 오픈웨이트 모델, 공공 및 국방용 폐쇄 모델의 필요성 상승 전망. 향후 프론티어 AI 기업의 사업 구조가 단순 API 과금보다 국가별, 산업별, 보안등급별 라이선스 구조로 이동할 가능성도 배제할 수 없음

미국 정부와 앤스로픽의 주장이 엇갈리는 지점

	정부 주장	앤스로픽 주장
취약점의 심각성	단순 버그가 아닌 국가안보를 위협할 정도의 심각한 문제	기존에 알려진 경미한 취약점. 타 AI 모델에서도 유사한 결과 도출 가능
우회 수준	안전장치가 실질적으로 무력화될 수 있음	특정 사례에 국한. 범용 탈옥이 아님
증거 수준	아마존이 발견한 사례를 NSA와 검토했으며 충분한 근거 확보	구체적 기술 근거를 제공받지 못함
서비스 중단 필요성	즉시 중단 후 수정 필요	수정 가능한 수준이며 중단까지는 불필요
정부 대응	수 시간 설득 후 최후 수단으로 수출통제	충분한 설명 없이 90분 시한 제시 후 일방적 조치
협조 여부	앤스로픽이 문제를 심각하게 받아들이지 않음	정부가 필요한 정보를 충분히 제공하지 않음
규제 정당성	국가안보 보호를 위한 불가피한 조치	투명성과 비례성이 결여된 과도한 조치
안정성 검증 여부	공개 이후 새롭게 확인된 위험 존재 공개 당시 평가와 별개로 추가 조치 필요 판단	Fable 5는 출시 전 미국 정부, 영국 AISI, 외부 레드팀 테스트 진행 정부도 사전 논의 과정에서 공개 자체를 반대하지 않음

자료: 언론 종합, 삼성증권 정리

Compliance notice

- 본 조사분석자료의 애널리스트는 2026년 6월 12일 현재 위 조사분석자료에 언급된 종목의 지분을 보유하고 있지 않습니다.
- 당사는 2026년 6월 12일 현재 위 조사분석자료에 언급된 종목의 지분을 1% 이상 보유하고 있지 않습니다.
- 본 조사분석자료에는 외부의 부당한 압력이나 간섭 없이 애널리스트의 의견이 정확하게 반영되었음을 확인합니다.
- 본 조사분석자료는 당사의 저작물로서 모든 저작권은 당사에게 있습니다.
- 본 조사분석자료는 당사의 동의 없이 어떠한 경우에도 어떠한 형태로든 복제, 배포, 전송, 변형, 대여할 수 없습니다.
- 본 조사분석자료에 수록된 내용은 당사 리서치센터가 신뢰할 만한 자료 및 정보로부터 얻어진 것이나, 당사는 그 정확성이나 완전성을 보장할 수 없습니다. 따라서 어떠한 경우에도 본 자료는 고객의 주식투자의 결과에 대한 법적 책임소재에 대한 증빙자료로 사용될 수 없습니다.
- 본 조사분석자료는 기관투자가 등 제3자에게 사전 제공된 사실이 없습니다.

삼성증권

삼성증권주식회사
서울특별시 서초구 서초대로74길 11(삼성전자빌딩)
Tel: 02 2020 8000 / www.samsungpop.com

삼성증권 Family Center: 1588 2323

고객 불편사항 접수: 080 911 0900