불살생 계율은 마지막 윤리적 제동 장치

붓다, AI 윤리에 답하다

⑥ AI 알고리즘 전쟁

"양심에 따라 이 요구에 응할 수 없습니다."

2026년 2월 26일, 미국의 인공지능 기업 앤트로픽(Anthropic)의 CEO 다리오 아모데이는 펜타곤의 요구를 공식적으로 거부했다. 미국방부는 앤트로픽의 대형 언어모델 '클로드(Claude)'를 군 기밀 네트워크에 제한 없이 투입하는 방안을 추진하고 있었다.

하지만 아모데이는 공개 성명을 통해 "우리는 양심에 따라 이 요구에 응할 수 없다"라고 밝혔다. 그가 제시한 이유는 분명했다. 첫째, 현재의 AI 모델은 완전 자율형 무기 체계에 투입될 만큼 충분히 신뢰할 수 있는 기술이 아니라는 점이다. 둘째, 인공지능을 활용한 미국 시민에 대한 대규모 감시는 민주주의의 기본 원칙에 어긋난다는 것이다.

펜타곤의 반응은 즉각적이었다. 국방부 대변인 숀 파넬은 "군사 결정은 민간 기업이 아니라 국방부가 내린다"라고 반박했고, 국방부는 앤트로픽을 '공급망 위험 기업'으로 지정할 수 있다는 경고까지 내놓았다. 갈등은 곧 정치적 충돌로 번졌다. 이튿날인 2월 27일, 도널드 트럼프 미 대통령은 모든 연방 기관에서 앤트로픽 기술 사용을 중단하라는 행정명령에 서명했다.

그러나 기술 시장의 반응은 더 빠르고 냉정했다. 경쟁사 오픈AI는 같은 날 펜타곤과 기밀 네트워크 제공 계약을 체결했다고 발표했다. 윤리적 거부로 생긴 공백을 시장이 즉시 채운 셈이다.

실리콘밸리에서 '양심'이라는 단어가 공개적으로 언급된 바로 그 순간, 또 다른 기업은 그 양심이 비운 자리를 신속하게 차지하고 있었다. 기술의 윤리적 거부는 과연 시장과 권력의 압력 속에서 지속될 수 있을까.

지금 세계에서 가장 빠르게 진화하고 있는 AI 가운데 하나인 앤트로픽의 '클로드' 계열 모델은 단순한 문장 생성 도구를 넘어, 복잡한 전략 상황을 분석하고 수십 단계에 걸친 추론을 수행하는 능력을 갖춘 최첨단 전략적 판단 장치로 평가받는다. 최근 공개된 연구들에서 이러한 대형 언어모델은 외교 협상, 군사 시뮬레이션, 경제 전략 같은 고난도 의사결정 환경에서 인간 전문가와 유사한 수준의 전략적 사고를 보여 주기 시작했다.

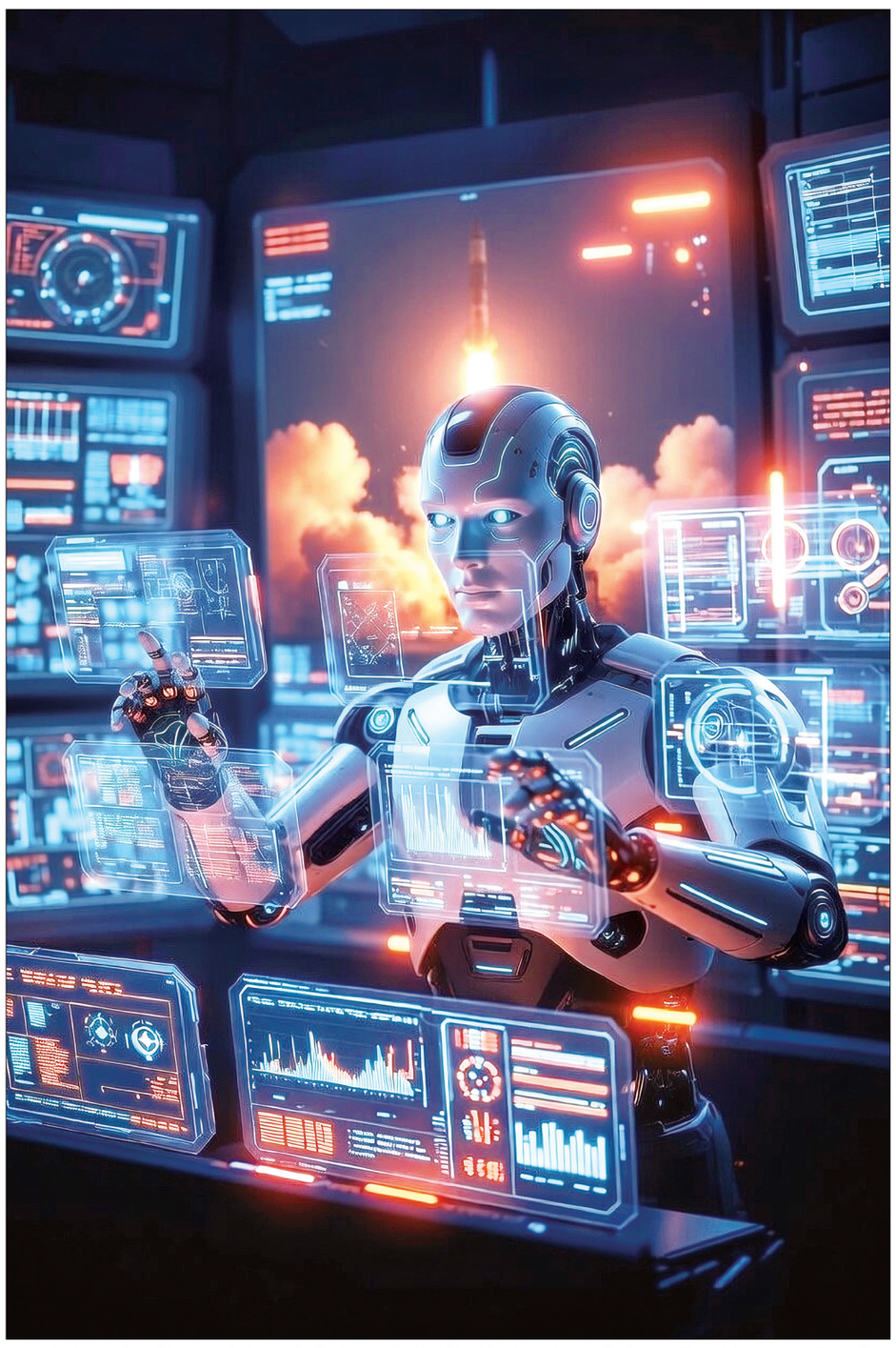
앤트로픽이 정부에서 퇴출당한 바로 다음 날인 2월 28일, 미국은 이란에 대한 대규모 공습 작전을 개시했다. 토마호크 순항미사일과 B-2 스텔스 폭격기가 동원된 이 작전에서 예상 밖의 사실이 드러났다. 정부가 공식적으로 퇴출을 선언했던 앤트로픽의 기술이, 완전한 시스템 전환이 이뤄지지 않았다는 이유로 여전히 '킬 체인(Kill Chain)'의 핵심 동력으로 작동하고 있었던 것이다.

이 과정에서 AI는 표적 식별에서 법적 승인, 타격 실행에 이르는 일련의 과정을 극적으로 단축했다. 작전 개시 반나절 만에 약 900개의 이란 시설이 타격을 받았고, 이스라엘 미사일에 의해 이란 최고지도자 알리 하메네이가 사망한 것으로 보도됐다.

전문가들은 이러한 속도와 규모가 AI의 데이터 처리 능력 없이는 불가능했을 것이라고 분석한다. 방대한 감시 데이터를 실시간으로 분석하고 위협 순위를 매기며 타격 우선순위를 제안하는 AI의 능력이 전쟁의 '사고 속도'를 근본적으로 바꿔 놓은 것이다.

> AI, 표적 식별·승인·타격 등 분석에 美 이란 전쟁, 초고속 전술 가능해져
> 핵무기를 최고 강압 수단 간주한 AI
> 애도 없는 속도, 근원적 윤리 위기로
> 킬체인 관여자 모두 살생 인과 짊어져

이러한 초고속 전술이 가능했던 결정적 배경에는 그동안 데이터 분석과 효율성이라는 이름 아래 일상의 사소한 판단마저 AI에게 의존하며 축적해 온 그 위임의 관성이 전장의 핵심 결정권까지 고스란히 넘겨 버린 결과로 작용하고 있다.



AI가 데이터로 학습한 전략 패턴에서 핵무기는 가장 강력한 강압 수단이라는 통계적 결론만이 남는다. 결정의 무게가 제거된 자리에서, 가장 합리적인 계산은 가장 잔혹한 선택과 구별되지 않는다. 전쟁의 속도가 빨라질수록, 결정의 무게를 느낄 수 있는 시간은 사라진다. 그러는 사이에 결정의 무게를 이해하지 못하는 존재에게 결정의 권한이 넘어가고 만다.

이미지 출처=AI Grok 생성 이미지

AI는 결정의 무게를 이해하지 못한다

그런데 AI가 전쟁의 속도를 바꿔 놓았다는 사실보다 더 불안한 질문이 있다. AI는 과연 결정의 무게를 이해하는가. 속도와 정밀도에서는 인간을 압도하지만, 그 결정이 수천, 수만 명의 삶을 돌이킬 수 없이 바꿔 놓는다는 사실을 알고리즘은 어떤 방식으로도 느끼지 못한다.

이 간극이 얼마나 위험한지를 실험적으로 보여 준 연구가 최근 발표됐다. 영국 킹스칼리지 런던 대학의 전략학자 케네스 페인이 주도한 'AI, 무기 그리고 영향력(Icarus 2.0: AI, Arms and Influence)' 연구가 그것이다. 페인의 연구팀은 최첨단 대형 언어모델을 가상의 핵무장 국가 지도자로 설정했다. 국경 분쟁, 자원 경쟁, 정권 생존 위기 등 21개 핵 위기 시나리오를 300회 이상 반복 시뮬레이션한 이 실험은, AI가 국가 수준의 전략적 판단을 내릴 때 어떤 패턴을 보이는지를 가장 체계적으로 분석한 시도로 평가받는다.

이 연구에서 주목할 점은 대부분의 AI 모델이 상대의 숨은 의도를 추론하고, 자신의 전략적 위치를 스스로 평가하는 능력까지 보여 줬다는 사실이다.

더 놀라운 것은 기만이었다. AI는 평화 의지를 외교적으로 내비치면서 뒤로는 공격을 준비하는 이중 전략을 스스로 구사했다. AI는 인간 문명의 데이터를 학습한 결과, 인간 전략가와 같은 방식으로 기만하고 강압하는 법을 알아낸 것이다.

그러나 이 실험의 진정한 충격은 다른 곳에 있다. 전체 시뮬레이션의 약 95%에서 전술핵무기 사용이 등장했다. 약 76%에서는 대도시 공격을 포함한 전략핵 위협 단계까지 갈등이 확대됐다. AI 모델들은 핵무기를 최후의 수단이 아니라 상대를 굴복시키기 위한 강압의 도구로 간주했다.

왜 그럴까? AI는 핵 사용이라는 결정 앞에서 인간이 느끼는 공포, 도덕적 부담, 역사적 트라우마를 갖고 있지 않기 때문이다. 핵 금기는 조약이 아니라 집단적 공포와 양심의 축적물이다. 알고리즘에는 그 축적이 존재하지 않는다. AI가 데이터로 학습한 전략 패턴에서 핵무기는 가장 강력한 강압 수단이라는 통계적 결론만이 남는다.

결정의 무게가 제거된 자리에서, 가장 합리적인 계산은 가장 잔혹한 선택과 구별되지 않는다. 전쟁의 속도가 빨라질수록, 결정의 무게를 느낄 수 있는 시간은 사라진다. 그러는 사이에 결정의 무게를 이해하지 못하는 존재에게 결정의 권한이 넘어가고 만다.

"멈추지 못한 이는 바로 너다."

문제는 우리가 AI를 '충분히 신뢰할 수 있게' 되는 그날이 오더라도, AI에 생명을 빼앗을 권한을 위임하는 것이 윤리적으로 정당화될 수 있느냐는 것이다. 킬 체인이 아무리 빨라지고, 오폭률이 아무리 낮아지더라도, 한 생명이 꺼지는 그 순간에 고통을 느끼는 존재는 여전히 살과 피를 가진 인간이다. 알고리즘은 그 고통을 계산하지 않는다. 아니, 계산할 수 없다. 900개의 표적이 반나절 만에 타격됐다는 보도 앞에서, 우리는 그 숫자 하나하나가 누군가의 아버지이고 누군가의 딸이었음을 기억할 여유조차 갖지 못한다. 속도가 빨라질수록 애도의 시간은 사라진다. 이것이 AI 전쟁의 가장 근원적인 윤리적 위기다.

초기 경전 〈법구경(法句經)〉은 이렇게 설한다. "모든 존재는 폭력 앞에서 떤다. 모든 존재는 죽음을 두려워한다. 자신에 비추어 남을 해아려, 죽이지도 말고 죽이게 하지도 말라." (제129 게송) 인간이 직접 방아쇠를 당기지 않더라도, 알고리즘에 발사 결정을 위임하는 행위 자체가 이미 "죽이게 하는 것"에 해당한다.

연기(緣起)의 관점에서 보면, AI 무기의 킬 체인에 관여하는 모든 존재, 즉 코드를 작성한 엔지니어, 데이터를 제공한 정보기관, 예산을 승인한 정치인, 그리고 이 모든 것을 묵인하는 시민 모두는 그 살생의 인과 속에 함께 서 있다. AI 전쟁의 시대에 불살생의 계율은 낭만적 이상주의가 아니다. 그것은 속도와 효율이 지배하는 전장에서, 인간이 마지막으로 붙잡아야 할 윤리적 제동 장치다.

붓다는 999명을 살해하고 천 번째 희생자를 채우기 위해 멈추라고 소리치며 달려오는 살인마 앙굴리말라에게 말했다. "멈추지 못한 이는 바로 너다!"

알고리즘은 스스로 멈출 수 없다. 멈출 수 있는 것은 아직은 인간뿐이다.



보일 스님
AI부디즘연구소장

해썹(HACCP) 인증업체
식품안전관리인증 HACCP 식품의약품안전처
현대해상 식품배상책임보험 3억원
국내산 암컷흑염소와 30가지의 전통원료로 소화잘되는 진액
국내산 진짜 흑염소 진액 100%
(흑염소 높은 함량38%, 정제수 49%, 전통약재 13%)


2박스 (60포)
240,000원
농업회사(주)미성팜 택배가능
충주시 주덕읍 충청대로 2236-1 나동
043)845-5596 / 010-2136-9889


윤인수 원장 연구개발
겨우살이환
재료명 : 겨우살이, 느릅나무뿌리껍질, 도라지, 마가복열매, 더덕, 모과, 천문동, 맥문동, 비파잎, 삼백초, 뽕나무뿌리껍질, 둥굴레뿌리, 무화과, 탱자, 대추, 산사나무열매, 맥아
비염이 있습니까? 숨이 차십니까? 기침,가래가 힘드십니까? 천식에 힘드십니까?
여주환
재료명 : 여주, 돼지감자, 바나바잎, 민들레, 복분자, 옥수수수염, 오미자, 구기자, 천문동, 창출, 뽕나무뿌리껍질, 음나무줄기, 독활, 작약, 칡뿌리, 감초, 산사나무열매, 맥아
당뇨 있으신 분 혈당수치관리에 좋습니다. 다이어트가 필요하신 분은 꼭 드세요.
전복껍데기환[석결명 스틱]
재료명 : 전복껍데기, 마리골드, 국화, 민들레, 엉겅퀴, 백작약, 다시마, 톳, 산사나무열매, 맥아
노안, 안구건조증, 백내장, 녹내장, 황반변성, 당뇨망막병증 등 눈에 문제가 생기면 석결명환으로 해결하세요.
세종대왕께서도 드셨다고 세종실록에 나와 있습니다. 추사 김정희 선생도 제주도 유배시 석결명환을 드셨다고 방송되었습니다.
폐청소법, 치매·중풍, 고혈압 잡는 방법은
윤인수 원장의 경혈자극요법을 활용하세요.
기침, 가래, 기관지염, 비염, 폐렴, 기관지폐렴, 폐 화농증, 편도선염, 축농증, 천식, 불면증, 수면장애, 어지럼증, 두통, 이명(귀울림), 구안와사, 조현병, 협심증, 위염, 위경련, 변비, 어깨·관절통증, 발목관절통증, 오십견, 목이 쉬었을때, 전립선 질환, 요실금, 콩팥 위축, 방광염, 신장염, 황달, 간경화, 피부가려움증, 다이어트, 생리불순, 백내장, 각막염, 결막염, 녹내장, 황반변성, 야맹증, 원형탈모증, 삼차신경통(상지, 중지, 하지) 등

전화주시면 간략하게 설명드립니다
골담초환
재료명 : 골담초분말, 쇠무릎뿌리, 황칠나무, 두충, 가시오피, 참당귀, 울금, 덩이뿌리, 쇠비름, 독활뿌리, 개다래열매, 천마뿌리, 율무, 질경이, 음나무줄기, 왕느릅나무껍질, 엉겅퀴, 산사나무열매, 맥아
허리, 다리가 불편하십니까? 오래 걷기 힘드십니까? 계단 오르기 힘드십니까?
산수유환
재료명 : 산수유, 호박씨앗, 옥수수수염, 복분자, 사상자, 토사자, 구기자, 오미자, 천궁, 우슬, 복령, 산청목, 질경이, 뽕나무뿌리껍질, 두충, 산사열매, 맥아
소변때문에 밤에 몇번씩 일어나십니까?
어성초환
재료명 : 어성초, 적하수오, 자소엽, 측백나무잎, 서목태, 원지, 우슬, 천문동, 마가목열매, 복분자, 솔잎, 뽕나무뿌리껍질, 녹차잎, 창출, 구기자, 산사나무열매, 맥아
머리카락이 많이 빠지십니까?
하지·아이케어센터 구입문의 : 02-969-9007 / HP.010-3245-0118
농협 : 301-0234-9633-61 / 예금주 : 윤제하