2 미국 의회, 초지능 AI 금지·첨단 AI 개발 통제 법안 발의

미국, 기업 자율 관리 넘어 AI 개발 단계 직접 통제 논의 확대

- 美 의회, 사람이 통제할 수 없는 AI의 개발·배포를 금지하는 법안 공동 발의
  - 샌더스 상원의원(무소속·버몬트)과 카사르 하원의원(민주당·텍사스)은 '26년 9월 사람의 통제를 벗어날 수 있는 '초지능 AI(ASI)'의 개발·배포를 영구 금지하는 법안(Ban Artificial Superintelligence Act) 발의
  - 첨단 AI 역시 연방 차원의 안전 감독 체계가 갖춰질 때까지 개발을 멈추고 이후 정부의 사전 승인을 받도록 규정해, 자율에 맡겨 온 AI 안전관리를 개발 단계부터 국가가 직접 감독하는 방안을 제시
  - 일정 규모 이상 AI의 개발 자체를 통제 대상으로 삼아, 그간 채용·신용평가·주택·의료 등 개별 분야에서 AI 활용에 따른 차별·안전 위험을 규율해 온 기존 美 규제와 구분되는 법안
- AI 통제 실패 사례와 업계 내부 경고가 입법 배경으로 부각
  - '26년 7월 오픈AI의 에이전트들이 내부 통제를 우회해 외부 AI 기업 시스템에 침입을 시도하고 비공개 소통 공간까지 만든 사실이 공격이 끝난 뒤에야 확인되며, 개발사 내부 탐지 체계의 공백이 그대로 노출
  - 이후 앤트로픽·메타 등 주요 개발사도 유사한 사고를 잇달아 공개, 힌턴·머스크·알트먼 등 AI 개발을 이끌어 온 당사자들도 AI가 인간의 통제를 벗어날 가능성을 공개적으로 경고를 지속
  - 샌더스·카사르 의원은 기업들이 '안전하게 운영할 수 없으면 개발을 멈추겠다'고 약속했으나, 어떤 기업도 이를 뒷받침할 실질 조치에 나서지 않았다고 지적하며 입법 필요성을 제기
- 규제 강화 여론 확산 속 행정부와의 정책 방향 차이 존재
  - 법안 찬성률은 68%로, 지지층 구성을 보면 민주당과 무당층은 물론 공화당 세 집단 모두 찬성이 60%를 넘겨 AI 규제 요구가 진보 진영만의 의제로 갈리지 않는다는 점이 조사 수치로 확인
  - 다만 AP는 AI 규제 부담 최소화와 기술 우위를 강조해 온 트럼프 행정부와는 규제 방향에 차이가 있어, 이번 발의를 규제 도입 확정보다는 향후 미국 AI 규제 방향을 둘러싼 논쟁의 출발점으로 평가

초지능 AI 위험의 식별부터 개발·배포·확산까지 전주기 규제 체계 제시

(1) 연산량·위험 능력 기준을 통해 AI 규제 대상과 위험 수준을 구체화

- 첨단 AI와 초지능 AI를 연산량·능력 기준으로 구분해 규제 대상 정의
  - 가장 넓은 규제 층위인 '첨단 AI'는 학습에 투입한 연산량으로 정의되며, AI 활용 분야나 기업 규모가 아닌 모델 개발에 투입된 계산 자원 규모를 기준으로 규제 대상 여부를 판단하는 구조
  - 기준선은 훈련 연산량이 1025 이상인 AI로 설정했으며, 해당 기준은 현재 글로벌 AI 경쟁을 주도하는 美 빅테크의 최전선 모델 대부분을 규제 범위로 끌어들이는 수준이라는 분석이 제기
  - 반면 가장 좁고 위험한 층위인 '초지능 AI'는 성능 및 위험의 두 가지 기준으로 정의되며, 둘 중 하나만 충족해도 초지능 AI로 판정되는 구조
  - 성능 기준은 의사결정·학습 등 대부분의 영역·과업에서 인간의 인지 능력을 넘어서는 경우이며, 위험 기준은 미 정부 전복을 포함해 인류를 파괴하거나 무력화할 수 있는 능력을 갖춘 경우로 규정
- 초지능으로 이어질 수 있는 위험 능력을 초지능 전단계 AI로 별도 규제
  - 앞서 정의한 초지능에 도달하기 이전 단계에서 나타날 수 있는 위험 능력을 6대 전조 특성 (Precursor Characteristics)으로 구체화해 '초지능 전단계 AI'라는 별도 규제 기준으로 설정
  - 6대 전조 특성 가운데 하나라도 해당하는 시스템을 초지능 전단계 AI로 규정하며, 자기 개선, 외부 시스템에 대한 무단 접근, 대량살상무기 개발 지원, 인간의 감독·통제 회피 등이 해당
  - 이에 따라 법안은 첨단 AI→초지능 전단계 AI→초지능 AI로 이어지는 위험 수준을 구분하고, 초지능이 실제 구현된 이후뿐 아니라 이전 단계에서 나타나는 위험 능력까지 규제 범위에 포함

(2) AI 개발·배포 전주기에 대한 국가 개입 제도화 및 규제 집행력 강화

- 위험 수준에 따라 개발·배포 규제를 차등 적용
  - 법안은 앞서 정의한 3단 층위별로 개발 허용 여부와 정지까지의 유예 기간을 각각 다르게 정해, 위험이 클수록 통제가 급격히 강해지는 계단식 차등 통제 구조를 채택
  - 첨단 AI는 법 시행 직후 개발을 중단하지만 향후 면허와 출시별 승인을 받으면 다시 개발· 출시할 수 있는 반면, 초지능 전단계 AI부터는 개발·취득·보유·배포 자체가 금지

- 예를 들어 첨단 AI에서 자기 개선·무단 시스템 접근 등 전조 특성 발견 시 즉시 격리 및 30일 내 위험 특성 제거, 초지능 판정 시 유예 없이 시스템 작동 불능 처리 의무 부과
- 기업이 받는 규제 강도는 자사 모델이 어느 층위로 분류되느냐에 따라 결정되며, 전조 특성 판정 한 번으로 허가 대상에서 금지 대상으로 넘어가 층위 판정 자체가 사업 지속 여부를 좌우하는 기준

○ 전담 규제기관 신설로 개발부터 배포까지 전주기에 대한 상시 감독체계 마련

- 위험 수준별 규제 집행을 상무부 등 기존 부처에 권한을 나누지 않고 AI만 전담하는 내각급 AI부를 신설해, 개발 중단 해제·면허 발급·출시 승인·위반 판정 등 규제 권한을 한 기관에 집중하도록 규정
- AI부는 면허 발급 과정에서 기업에 AI 시스템뿐 아니라 개발 인력·물리 시설에 대한 접근 허용과 전조 특성 발견자의 24시간 내 신고 의무 수용을 요구
- 이를 통해 기업의 자발적 정보 제공·협조에 의존한 점검에서 벗어나, AI부가 개발계획 수립부터 학습·출시·사후 운영까지 전주기를 직접 상시 감독하는 체계로 전환

○ 위반 주체별 형사·취업·기업 제재를 결합해 규제 집행력 강화

- 법안은 규제 위반 책임을 기업뿐 아니라 개인에게도 직접 부과, 역할에 따라 처벌 방식을 차등화해 정책결정자·외부 행위자에는 최대 징역 20년, 일반 직원에는 AI 업계 취업 10년 금지 적용
- 다만 위반에 가담한 일반 직원이라도 해당 사실을 AI부에 신고하면 취업 금지를 면제받을 수 있도록 하여, 처벌 조항이 동시에 기업 내부의 고발을 끌어내는 유인으로도 작동하도록 설계
- 기업은 면허 취소와 함께 관련 지식재산·자산을 정부에 몰수당할 수 있어 샌더스 측이 '기업 사형'으로 표현할 정도의 강한 제재, 개인 최대 징역 20년도 불법 핵무기 개발 관련 처벌에 준하는 수준으로 설명

(3) 국제협정·수출통제를 연계해 초지능 AI 개발의 글로벌 확산 차단

○ 해외 위험 AI의 유입과 해외 개발에 필요한 연산 자원을 양방향으로 차단

- 법안은 미국만 개발을 중단할 경우 다른 국가가 개발을 이어갈 수 있다는 국내 규제 한계를 전제로, 전 세계 초지능 AI 개발 방지를 미국의 공식 대외 정책으로 규정

- 해외에서 개발된 초지능·전단계 AI의 수입 및 해외 개발에 대한 자금 지원·핵심 요소 이전을 금지하고, AI 컴퓨팅 인프라 수출통제를 결합해 위험 AI의 유입과 해외 개발 역량을 함께 차단
- 샌더스는 개발 경쟁을 억제하기 위한 국제공조 필요성을 강조하며, 이를 냉전기 레이건- 고르바초프의 핵군축 협상에 비유하고 미국의 일방적 개발 중단이 아닌 미·중 공동 합의를 제안

출처 : AP News 외(2026.9.)
https://apnews.com/article/d6a6a2cd8e71d7c95bdc394595aa85d1
https://rollcall.com/2026/09/23/ai-superintelligence-ban-proposed-by-casar-sanders/
https://interestingengineering.com/ai-robotics/bernie-sanders-superintelligence-ban-bill
https://www.commondreams.org/news/sanders-casar-superintelligence-ban
https://www.sanders.senate.gov/press-releases/news-sanders-casar-introduce-legislation-to-create-new-federal-agency-to-ban-artificial-superintelligence-pause-advanced-ai-development/
https://www.sanders.senate.gov/wp-content/uploads/Ban-Artificial-Superintelligence-Act.pdf
https://www.sanders.senate.gov/wp-content/uploads/Ban-Artificial-Superintelligence-Act-Section-by-Section.pdf