6 CSIS, 미국 주정부·EU·산업계의 AI 규제 프레임워크 분석 보고서 발표

미국 전략국제연구센터(CSIS)는 미국 주정부들의 AI 규제 법안과 EU AI법(AI Act), 산업계의 자발적 AI 안전 프레임워크를 비교·분석한 보고서*를 발표('26.8)

* Toward a Federal Framework: Lessons from State and International Frontier AI Regulation

○ **(개요)** 주정부 차원에서 추진되는 프론티어 AI 규제 법안들과 EU 및 산업계의 AI 규제 프레임워크를 비교해 미국 연방 차원의 AI 규제 입법을 위한 제언을 제시

- 9가지 프레임워크들이 AI 모델 규제 방법론에 있어 여러 공통 요소를 가지고 있다고 분석하고, 이러한 공통 기반이 글로벌 기준선이 되어가고 있다고 설명
- 다만 규제 대상·범위·강도 등에서 여전히 상당한 불일치도 존재한다고 지적
- 현재 미국 연방 차원의 AI 규제 프레임워크가 부재한 가운데, 주정부들의 AI 규제 실험을 '규제 파편화'나 '통일된 프레임워크를 가로막는 방해물'로 간주하기보다 연방정부의 AI 입법 설계에 활용할 것을 제언

〈 CSIS가 분석한 9가지 AI 규제 프레임워크 개요 〉

구분	주요 내용
캘리포니아주 S.B. 1047	• 대규모 AI 모델 개발사뿐만 아니라 컴퓨팅 클러스터 운영자 및 제3자 감사인까지 규제 대상에 포함시키는 등 매우 강력한 AI 규제 법안 • 보안 조치, 위험 평가, 사고 보고, 고객확인(KYC) 요건은 물론 일정 수준 이상의 위험 발생 시 모델 작동을 강제 종료하는 '킬 스위치' 조항도 포함 • 주지사의 거부권 행사로 폐기('24년 9월)
캘리포니아주 S.B. 53	• S.B. 1047보다 훨씬 완화된 개발자 중심 AI 모델 규제법 • AI 개발사에 안전 및 보안 프로토콜(SSP*) 공개 및 중대 사고에 대한 신고 등을 의무화 * safety and security protocol: 기업의 거버넌스 및 규정 준수 전략을 설명하는 문서로, 대개 위험 평가, 위험 완화·보안 조치 등을 포함 • '25년 제정되어 '26년 1월부터 시행 중
뉴욕주 RAISE 법 (Responsible AI Safety and Education Act)	• 대형 프론티어 AI 모델 개발사에 SSP 공개와 사고 보고 의무를 부과 • '25년 12월 주지사 서명을 통해 법률로 제정
일리노이주 S.B. 315	• 대형 프런티어 AI 개발사에 대해 독립된 제3자를 통한 연례 안전 감사를 의무화 • '26년 7월 주지사 서명을 통해 법률로 제정
미시간주 H.B. 4668	• AI 모델에 대한 안전·보안 요건, 위험 평가, 투명성 의무 등을 규정 • '25년 6월 발의되었으며 현재 의회 계류 중
매사추세츠주 S. 2630	• AI 모델에 대해 안전·보안 조치, 위험 평가·완화, 사고 보고, 내부 고발자 보호 등을 요구 • '25년 10월 발의되었으며 현재 의회 계류 중

구분	주요 내용
EU AI법	• 세계 최초의 포괄적인 AI 규제법 • AI가 미치는 위험 수준에 따라 차등적인 규제 적용 • 동 법에 명시된 범용 AI(GPAI) 모델 관련 의무 이행을 지원하기 위해 EU가 채택한 '범용AI 실천강령'('25.8.1.) 역시 법적 구속력을 보유 • '25년 7월 공표된 후 범용AI 실천강령 적용이 시작('25.8.2.~)되었고, 시스템적 위험을 가진 범용AI 관련 집행·제재도 본격 시행('26.8.2.~)
서울 프론티어 AI 안전 서약	• '24년 5월 AI 서울 정상회의에서 발표된 AI 기업 간 자발적 안전 서약 • AI 기업들은 프런티어 AI 모델의 위험 식별·평가, 위험 완화 조치 및 SSP 마련, 일정한 위험 임계치에 도달 시 책임 있는 조치 등을 약속
앤트로픽의 프론티어 모델 투명성 프레임워크	• '25년 7월 앤트로픽이 제안한 자발적 AI 안전 관리 프레임워크 • 초대형 모델과 개발사에 한정해 안전 관행 및 위험관리의 투명성을 높이는 비교적 가벼운 안전 프레임워크

○ **(공통 특징)** CSIS가 분석한 9개 AI 규제 프레임워크는 ▲안전 프레임워크 ▲투명성 ▲위험 평가 ▲위험 완화라는 4가지 조치들을 공통적으로 포함

〈 9가지 AI 규제 프레임워크 비교·분석 결과 - 공통 요소 〉

구분	주요 내용
안전 프레임워크	• 안전 및 보안 프로토콜(SSP) 또는 이와 유사하게 일련의 심각한 위해에 대한 리스크 완화 전략, 보안 조치 등을 명시하는 프레임워크를 구현하고 정기적으로 업데이트하도록 의무 부여
투명성	• SSP를 일반에 공개해야 하는 의무 부여(필요시에는 상업적으로 민감한 일부 내용을 삭제해 공개)
위험 평가	• 잠재적으로 위험한 AI 모델의 기능을 식별·평가하도록 요구
위험 완화	• 재앙적인 결과를 예방하거나 줄이기 위한 조치를 마련하도록 요구

- 전 세계는 아직 프론티어 AI 규제 모델에 합의하지 못했으나 이러한 공통 조치는 AI를 책임 있게 감독하는 것에 대한 최소의 공통 기반을 형성
- 또한 CSIS가 분석한 프레임워크들은 앤트로픽을 제외하면 모두 모델 가중치·시스템에 대한 무단 액세스 방지와 같은 보안 조치를 요구
- 내부 고발자 보호를 위한 조치 역시 RAISE와 서울 서약을 제외한 대부분의 프레임워크에 포함
- '사고 보고 요건'도 다수의 프레임워크에 포함되었는데*, 보고 시한은 '72시간 이내'부터(RAISE에 해당) 사고의 심각성에 따라 보고 시한을 단계별로 적용하는 방법(캘리포니아 S.B. 53 및 EU의 범용 AI 실천 강령)까지 다양

* 서울 서약과 앤트로픽의 프레임워크에는 사고 보고 요건이 부재

● **(차이점)** 9개 AI 규제 프레임워크는 ▲규제 대상(가치사슬) 범위 ▲규제 모델 규모 ▲다루는 위험의 범주 ▲'재앙적 위험'의 정의 측면에서 차이가 존재

- 전반적으로 EU의 AI 규제는 현재 미국의 그 어떤 주 법안보다 더 많은 모델, 더 많은 위험, 더욱 광범위한 피해를 포괄

〈 9가지 AI 규제 프레임워크 비교·분석 - 차이점 〉

구분	주요 내용
규제 대상(가치사슬 내 행위자) 범위	• 프레임워크별로 의무를 이행해야 하는 AI 운영자(각 가치사슬 단계 및 관련 행위자)의 유형은 다양 • 캘리포니아주 S.B. 1047(현재는 폐기됨)은 AI 개발사는 물론이고 컴퓨팅 제공업체와 제3자 감사인에도 특정 의무와 규제를 부과
규제 대상 모델의 규모	• 미국 주정부 법안들은 바이든 행정부의 대통령 행정명령 14110호에서 영감을 받아 마련되어, AI 모델의 성능을 FLOP* 단위에서 구분해 높은 훈련 임계값과 규제 대상 기업을 정의하는 경우가 다수 * Floating Point Operations Per Second: 1초당 처리 가능한 부동 소수점 연산 횟수를 의미 • 반면 EU는 더 낮은 임계값을 설정하고 있으며, 소규모 모델을 시스템적 위험을 초래하는 모델로 직접 지정할 수 있도록 허용
위험의 범주	• 미국 주정부 법안들은 모두 화학·생물학·방사능·핵(CBRN) 위험, 인간의 통제가 거의 불가능한 심각한 범죄, 사이버 공격을 다루고 있으며, 대부분 AI에 대한 통제 상실도 위험의 범주에 추가 • EU는 여기서 더 나아가 대규모 조작이나 인권침해와 같은 사회적 위험도 위험의 범주에 추가
재앙적 위험에 대한 정의	• 미국 주정부 법안들은 모두 높은 인명 피해(사망자 50~100명 이상) 또는 경제적 손실(피해액 10억 달러 초과)을 기준으로 재앙적 위험을 정의 • 반면 EU는 인프라 붕괴, 인권침해, 환경적 피해까지 포함해 재앙적 위험을 정의하는 광범위한 접근방식을 채택

● **(연방 AI 정책에 대한 영향)** 연방정부는 주정부의 입법 실험을 억누르기보다 그 성과와 한계를 참고해 혁신을 과도하게 저해하지 않으면서도 통일된 연방 AI 규제 프레임워크를 마련하는 노력이 필요

- 지난 수년 동안 미국의 각 주는 수백 건의 AI 관련 법안을 발의하면서 AI 규제에서 중요한 시험대 역할을 수행해왔음
- 이러한 경험은 연방정부 차원의 AI 규제 접근방식 마련을 위한 유익한 정보로 활용 가능
- 따라서 주정부의 AI 입법 시도는 연방정부가 규제 파편화를 막고 기업의 규제 부담을 줄이는 연방 차원의 AI 입법을 가속화하는 데 중요한 청사진을 제공

출처 : 미국 과학기술정책실(OSTP) (2026.8.3.)
https://www.csis.org/analysis/toward-federal-framework-lessons-state-and-international-frontier-ai-regulation