



기정학 시대의 글로벌 AI 거버넌스 현안과 시사점¹⁾

- ➔ AI 기술의 급속한 발전과 AI 이용이 보편적 확산으로 인해 이에 대한 기준, 표준화, 국제규범에 대한 글로벌 거버넌스 논의가 본격화
 - AI 개발 및 혁신 차원의 개별 국가 차원의 전략·정책 수립을 중심으로 논의 되었으나, UN, OECD 등 기존 글로벌 AI 거버넌스 논의는 AI 이용 시 윤리적 기준과 원칙을 제시하는 수준
 - 글로벌 AI 거버넌스는 거버넌스의 하위 부분을 구성하는 협의체 및 국제기구 신설, 참여 행위자 구성, 핵심의제 및 국제규범 창설 등이 주요 현안
 - 미국, 중국, EU 등 기술 강국들은 국가전략 차원의 정책, 법제 정비를 적극적으로 추진함은 물론 자국의 정책 및 규제 방식을 지역 및 다자화를 통해 확산
- ➔ 미·중 및 주요국 간 기술 패권 경쟁은 AI 거버넌스 논의와 밀접히 연계되어 있으며 기술 강국들은 자국의 기술혁신을 촉진하면서도 경쟁국의 기술혁신은 제한하는 수단으로 표준 설정과 기술 통제를 강화하는 경향
 - 표준 및 규범 설정 문제는 신기술이 보편적으로 확산될 경우에 대두되는 이슈로, 신기술의 이중용도 사용 등 양면적 특성*에 기인
 - * 신기술의 개발·이용에 따른 편익(효용)과 안보·군사 분야 적용에 따른 파괴적 위협
 - 글로벌 AI 거버넌스는 상업적 활용과 군사적 이용을 명확히 구분하여 논의되었으나, 최근 논의 방향은 보다 복잡하게 전개되는 양상
 - UN에서의 논의 역시 상업적 이용 시 안전성, 국제 안보 맥락에서의 규범화, 자율 무기 체계의 대한 인간의 통제 등을 중심으로 거버넌스 및 규범 논의가 진행
- ➔ 유엔에서의 거버넌스 논의와 별개로 EU 등 지역적 차원의 규범 창설과 유엔 외 신규 논의체, 그리고 개별 국가가 이니셔티브를 주도하려는 다자 협의체 등 전반적으로 거버넌스 간 경쟁 역시 심화
 - 전통적으로 국제규범 창설을 주도해 온 EU는 개인정보보호법(GDPR), 인공지능법(AI Act) 등을 통해 미·중이 주도하는 AI 경쟁에 대응
 - 특히, 트럼프 행정부가 다자주의에 대한 미온적인 태도를 보이는 상황에서 중국은 '26년 7월 WAICO(세계인공지능협력기구)를 발족하여 AI 거버넌스를 주도

1) 세종연구소 유준구 선임연구위원(jkoo88@sejong.org)

본고는 저자의 개인적인 견해이며 과학기술정보통신부와 KISTEP의 공식적인 의견이 아닙니다.

- ➔ AI의 기술적 특성으로 인해 여타 신기술과 결합하여 새로운 안보 위협으로 부상하고 있으며 이는 기술 및 수출통제 관점에서 접근
- 미·중 간 AI 경쟁은 기술 경쟁은 물론 국제질서의 재편성과 연계되어 있으며 기정학(techno-geopolitics)의 변화를 초래
- 트럼프 행정부는 일련의 AI 관련 행정명령을 통해 경제 안보 공급망 강화 차원에서 AI 수출통제 및 기술통제 정책을 공세적으로 추진
- AI는 안보, 경제에 대한 파급효과가 상당하고 국가, 기업, 개인에 전 영역에 침투하고 있는 상황에서 AI 기술통제 정책의 경우, 경쟁국인 중국은 물론 EU, 일본, 한국 등 주요 기술 선진국에 미치는 파급력은 심대

1 글로벌 AI 거버넌스 논의 동향과 쟁점

가. 다자 협의체에서의 논의 현황

- ➔ 최근 글로벌 안보, 경제 전반에서 AI의 발달이 가속화되면서 AI의 잠재적 위험 및 위협에 국제사회가 공동으로 대응하는 한편, 보다 신뢰성 있고 책임 있는 AI의 활용을 촉진하기 위한 거버넌스 필요성 대두
- UN, OECD, EU, G7/G20 등 여러 글로벌 지역 협의체에서 AI 거버넌스 논의가 빠르게 진행되고 있으나 초기 수준임과 동시에 분절적·파편적인 한계
 - 기존 다자 협의체에서는 주로 AI 활용 시 발생할 수 있는 위험 요인들을 제거·감소 하려는 윤리적, 규범적 가이드라인을 제시하면서 AI 위험관리 프로세스를 구축하는 것이 주된 현안
- UN의 경우 '24년 9월 '미래정상회의(UN Summit for the Future)'를 개최, 'Global Digital Compact'를 채택하여 포괄적인 AI 거버넌스 논의를 시작
 - 구테레스 UN 사무총장은 '21년 '공동아젠다보고서(Common Agenda report)'를 통해 AI 거버넌스 구축의 시급성을 제기하면서 UN 차원의 국제기구 설립을 제안
 - 이후 AI 거버넌스 논의를 위한 UN 차원의 국제기구 설립을 제안하였고 '23년 이에 대한 연구를 위해 'AI 고위급 자문기구(HLAB-AI)'를 신설
 - AI 자문기구 중간 및 최종 보고서에서는 기존 다양한 유엔기구의 유사성을 검토하여 향후 신설될 AI 국제기구 신설을 위한 방향성을 제시
 - 또한 거버넌스 모델의 적합한 예시로 국제기구로 IAEA(국제원자력기구), IPCC(기후변화에 관한 정부 간 협의체) 등을 제안



※ IAEA와 IPCC가 제안된 주된 이유는 AI 위험성에 대한 검증, 모니터링 및 위험평가 필요성에 기인하는데 IAEA 모델이 우라늄 개발, 원심분리기, 농축 재처리 등이 데이터, AI 칩, AI 모델 등이 유사하다는 점에 착안

- 한편, 미·중은 물론 AI 기술 강국 간 AI 글로벌 거버넌스 논의를 위한 국제기구 신설 및 프레임워크 경쟁이 심화됨에 따라, '24년 미·중은 각기 유엔총회의 결의안(A/RES/78/265, A/RES/78/311)을 제안하여 채택
 - 주목할 점은 AI 글로벌 거버넌스 경쟁이 본격화되고 있음에도 불구하고 미·중이 각기 제안한 두 개의 결의안이 미·중 간 상호 지지를 통해 채택
 - 이는 미·중 간 AI 경쟁은 상호 제한적 성격이 현저함에도 불구하고 과거 미·소 간 핵 경쟁처럼 AI 전략대화를 병행하여 미·중이 글로벌 AI 경쟁을 주도하려는 의도
 - 상기 두 개의 결의안은 각각 AI의 책임 있는 이용 차원에서 위험평가와 검증 체계를 중시하는 미국과, AI의 지속가능한 이용, 기술격차 해소 및 역량 강화를 지적하며 Global South를 주도하려는 중국의 입장과 상이

➔ 글로벌 AI 거버넌스가 국제협의체, 기구별 다층적·분업적 체제로 진행되는 가운데, UN의 AI 거버넌스에 대한 논의도 군사 분야로까지 확대

- 제79차 유엔총회 제1위원회에서 우리나라와 네덜란드가 공동으로 상정한 “군사 분야 인공지능과 국제 평화 및 안보에 대한 합의” 결의안*이 유엔 회원국들의 압도적 지지로 채택

* UNIDIR ('24.11.6.) Artificial Intelligence in the Military Domain and its Implications for International Peace and Security

- 동 결의안은 유엔 내에서 채택된 최초의 군사 분야 AI 결의안
 - ※ 민간 분야에서는 '24년 3월과 7월 미국과 중국이 각각 주도한 유엔총회 결의안이 총의로 채택된 바 있음

〈 UN AI 결의안 주요 내용 〉

-
- 군사 분야 AI에 국제법 적용 확인
 - 군사 분야 AI 도입의 혜택 극대화 및 위험 최소화를 위한 노력 필요
 - 개도국-선진국 간 AI 격차 해소 및 역량 강화 강조
 - 국제기구, 학계, 시민사회 등 다중 이해관계자 접근법의 중요성 강조
 - 각 국가, 국제기구, 학계 등의 의견 수렴을 위한 유엔 사무총장 보고서 요청
-

- UN 외 체제에서도 한국과 네덜란드는 「인공지능의 책임있는 군사적 이용에 관한 고위급회의(REAIM)」 회의를 공동으로 개최하는 등 군사 분야 AI 규범 형성을 위한 국제사회의 노력을 주도
 - REAIM은 AI의 군사적 이용에 대한 최초의 거버넌스 논의로 '27년 제4차 회의(케냐)가 예정되어 있고 향후 UN 제1위원회와 연계 여부가 주목

- 기존 AI의 군사적 이용 규제 논의는 CCW(특정 재래식무기 금지협약) 산하 LAWS(치명적 자율무기체계) GGE(정부 전문가 그룹)를 중심으로 진행 중이나, 주요 진영 간 입장 차가 첨예
 ※ '13년 UN 차원의 논의 개시 이후 비공식 전문가 회의('14~'16년)를 거쳐 '16년 GGE가 공식 출범하였으며, '18년 자발적 행동 규범(11대 기본원칙)을 채택
- AI의 군사적 이용 규제는 점차 무기체계는 물론 군사적 이용 전반으로 확산 되는 추세
- 미국은 특히 “AI 및 자동 시스템의 책임 있는 군사적 이용에 관한 정치적 선언”을 발표하여 유사 입장인 국가들의 참여를 독려

〈 글로벌 AI 거버넌스 관련 규범 논의 기제 〉

Competing Platform: Segmented but Mainstreaming	
기존 UN AI 결의안 (안보 영역)	• UN, ('24.3) 비군사영역, Seizing the opportunities of safe, secure and trustworthy AI systems for sustainable development (A/RES/78/265) • UN, ('24.7) 비군사영역, Enhancing international cooperation on capacity-building of artificial intelligence (A/RES/78/311) • UN, ('24.12) 군사-안보영역, Artificial intelligence in the military domain and its implications for international peace and security (A/RES/79/239)
UN 내 신규 조직 (상업 영역)	• AI 고위급 자문기구(High-Level Advisory Body on AI) 등
UN 외 조직 (군사-안보 영역)	• REAIM 등 ※ Political Declaration on Responsible Military Use of AI and Autonomy ('23.02)
기타 AI 조직 및 거버넌스	• (기구) UN 산하기구: UNESCO, ITU 등 • (선언) AI Safety Summit: 블레츨리 선언 등 • (가이드라인) OECD 2019 AI 원칙, G20 AI 가이드라인 등 • (입법) EU AI 법(AI Act) 및 규정 등 • (협약) 인공지능과 인권, 민주주의, 법치주의에 관한 유럽평의회 프레임워크 협약 (Framework Convention on AI, Democracy, Human Rights, and Rule of Law) 등

출처 : 저자 작성

나. 논의의 주요 의제

- ➔ (원칙과 범위) 글로벌 AI 거버넌스 논의의 기본 방향은 AI의 안전하고 책임 있는 이용을 위한 원칙을 제시함과 동시에 인권, 투명성, 책임성, 편향성 완화 등을 강조
- 상기 원칙 및 기준에 대해 서방과 글로벌사우스 간 근본적 인식의 괴리가 있으며, 중·러 및 글로벌사우스 진영은 “책임 있는”이라는 담론에 AI 이용에 자의적인 통제 의도가 있다고 비판
- 미국의 경우 안전하고 책임 있는 AI 이용의 모든 단계*에서 인권, 국제법, 개인정보보호, 지속 개발, 책임성 등을 강조하고 있는데 이는 필연적으로 위험 평가 및 검증체계를 수반



- * 사전설계(predesign)에서 해제(decommissioning) 단계까지 이르는 AI의 생애주기
- 중국은 글로벌사우스 차원에서 AI 기술격차 해소와 역량 강화를 부각시키면서 거버넌스 구축 과정에서 다수 국가의 지지를 견인
- 기존 글로벌 AI 거버넌스 논의 범위는 상업적 이용과 군사적 이용을 구분하여 진행
 - 다만, 상업적 이용에 적용되는 책임성, 투명성과 같은 윤리적 요소 및 주요 규범 원칙 등은 군사적 이용에서도 유사하게 적용
 - 문제는 AI 기술의 이중용도성으로 인해 상업, 군사적 이용의 구분이 모호하고 기술혁신 차원에서도 상업적, 군사적 이용을 동시에(spin-up) 고려하거나 상업적 개발의 즉각적 안보화(spin-off)가 빠르게 진행
 - 향후 AI 거버넌스 범위가 상업적, 군사적 이용으로 이원화된 것이 국제안보 맥락에서 포괄적으로 범주화될 가능성이 유력

➔ (군사적 이용) 기존 논의에서는 무기체계를 중심으로 CCW 체제 하에서의 기본 원칙을 제시해왔으나, 군사 분야에서의 AI 이용이 확대됨에 따라 무기체계를 포함한 군사분야 전반을 포괄하는 논의가 시작

- NATO는 AI의 군사적 이용 가능 분야를 C4ISR*를 명시한 바, AI의 군사적 이용의 범위는 상당히 광범위
 - * Command, Control, Communication, Computer, Intelligence, Surveillance, and Reconnaissance
 - 네덜란드와 한국이 순차적으로 개최한 'AI의 책임 있는 군사적 이용에 관한 고위급회의(REAIM)'는 이러한 추세를 반영
- LAWS에서는 핵 레짐(Nuclear Regime)과 같이 미·중·러의 이해가 일치하는 측면이 있지만, 포괄적인 군사 분야 논의의 경우 서방과 중·러 간 이해가 충돌
 - 비교적 적극적 논의를 주장하는 서방과 소극적인 중·러가 대립
- 자율무기체계의 규범 논의는 구속력 있는 조약 창설에 대해 유보적인 AI 기술 선진국과 창설을 적극 주장하는 기술 후발국 간 대립이 기본적 구도

다. 기술의 구성 요소

➔ (데이터 보호 및 이동) AI 관련 데이터 거버넌스의 경우 상대적으로 여타 쟁점보다 지속적인 논의가 있었던 쟁점

- △데이터의 자유로운 이동과 활용을 강조하는 미국, △주권에 따른 데이터 통제를 중시하는 중국과 개도국, △인권 및 개인정보보호 차원의 규범을 선도하는 유럽 간 입장이 상이

- 데이터 수집, 보관, 처리 및 이전에 관해 글로벌 차원에서 합의된 규범은 부재하며, 현 글로벌 데이터 규범은 상이한 양·다자간 프레임워크, 통상 규범, 국내 법률 등으로 파편화·분절화
 - 상이한 접근 방식을 추구하는 미국, 중국, EU는 글로벌 데이터 규범에 대한 영향력을 행사하며 자국 법·규범을 다자화 시도
 - APEC은 회원국 간 자유롭고 안전한 개인정보 이동을 지원하기 위해 기업의 개인정보 보호 체계를 평가하여 인정하는 글로벌 인증제도인 ‘국경 간 개인정보 보호규정(Cross Border Privacy Rules, CBPR)’을 시행(’23.8)
 - APEC CBPR은 참여국 기업을 위한 자발적 인증제도로, 현재까지 우리나라(8개사)와 미국(46개사), 싱가포르(11개사), 일본(5개사) 등 4개국에서 공공 및 민간기관이 인증을 부여
- 각국 정부는 국경 간 데이터 이동에 대한 규제를 채택하고 발전시키고 있으며, 국경 간 데이터 이동에 조건을 부여하거나 데이터의 현지 보관을 요구
 - 개인정보의 무분별한 국외 이동을 통제하기 위해 각국은 개인정보보호 법률을 통해 유사하면서도 조금씩 상이한 규범을 발전시켜 왔으며, 이는 AI 학습 등에 직접적인 영향
 - EU의 GDPR에서 규정한 적정성 인정 제도는 개인정보가 이전되는 국가의 정보보호 수준이 국내법과 비교하여 동등하거나 그 이상의 수준을 부여하는 것으로 인정하는 제도로, 적정성이 인정된 국가에는 개인정보를 자유롭게 이전
 - EU의 GDPR에서 처음 도입된 적정성 인정은 타 국가로 확대되는 추세로 영국을 포함해 일본, 호주 등에서 시행

⇒ (컴퓨팅) AI 컴퓨팅 거버넌스 논의는 최근에 부각되고 있으며 현재 미·중 간 이해가 가장 첨예하게 대립되는 영역으로, 경제안보 및 글로벌 공급망 논의와 긴밀히 연계

- 미·중 AI 패권 경쟁에서 데이터와 알고리즘과 달리 미국과 동맹국이 우월적인 지위를 누리고 있는 영역으로 미·중 국가전략의 핵심 의제로 접근
 - ※ AI 시스템 구성 요소로 컴퓨팅 분야는 데이터나 알고리즘과 달리 ‘경쟁적인 유형재화(rivalrous tangible good)’로 상당히 배제적인 특성
- 글로벌 AI 거버넌스 차원에서는 기술 및 수출통제 적용, 공급망 안전성, 기술 표준화 등이 주요 쟁점으로 미국은 자국 정책 및 법제의 다자화 작업을 적극적으로 추진
- 중국은 미국의 기술 통제에 대한 대응 전략을 강화하면서 미국의 글로벌 거버넌스 리더십이 소극적인 상황을 이용하여 거버넌스 주도권 확보에 적극적

- ➔ **(알고리즘)** 알고리즘/모델 거버넌스의 경우 개방형 모델과 폐쇄형 모델 경쟁이 주요 쟁점으로, 상대적으로 미국은 폐쇄형 모델, 중국은 개방형 전략을 추진
- 개방형과 폐쇄형 간 경쟁은 최근 ChatGPT와 DeepSeek의 경쟁을 통해 확인할 수 있으며, 중국과 같은 후발주자인 경우, 하이브리드 혹은 개방형 모델 선호
 - 민간 빅테크 기업이 주도하고 이를 국가 정책을 통해 뒷받침되고 있으며 개별 기업, 국가의 거버넌스가 글로벌 차원으로 확산되는 상황
 - 생성형 AI가 높은 부가가치를 창출할 것으로 예상되고 AI 기반 위협이 현실화됨에 따라 기존에 오픈소스를 기반으로 활성화되었던 AI 기술들이 점차 비공개화
 - 알고리즘의 안정성, 투명성 강화 수준 및 적정성 관련 논의가 글로벌 차원에서 국가 진영은 물론 빅테크 기업 간에도 이견이 상존
 - 이는 AI의 설명성 부재로 인해, 오작동이나 의도치 않은 임무 수행 등의 문제를 완화하기 위해 설명 가능성(white box)을 제고하는 일련의 규제 도입과 연계

2 글로벌 AI 국제규범 창설

- ➔ AI 시스템과 기술은 사이버, 우주와 같이 도메인이나 분야(sector)라기 보다는 도구적(tool) 성격이 강하며, 국제규범의 형성도 일정 정도 범주화되지만, 다른 도메인 및 분야로 연계·파편화되어 형성
- 사이버·우주 분야의 경우 도메인이나 분야의 관점에서 상대적으로 포괄적이며 범주화를 통해 국제규범이 형성되는 반면, AI는 상당히 분절적·파편적으로 규범화되는 경향이 현저
 - 범주화의 경우 국제 안보 맥락에서 국제 규범화 작업이 이루어지고 있으며 최근 안전 문제를 포괄
 - 국제법 창설이 상당한 노력과 시간이 소요되는바, 기존 국제법에서 AI 요소를 포함하는 방향으로 국제규범 형성
 - 사이버·우주·바이오·해양 등 기타 분야에서 AI 기술이 연계되는 위험 및 위협 요소를 반영하는 방식으로 국제규범 형성
 - 신기술로서 AI 관련 국제법의 적용 문제는 기존 국제법 분야 중 유엔헌장, 국제인도법 및 인권법, 국제형사법 분야가 주로 논의되고 있으며 기존 협의체의 결과물이나 메뉴얼 형태로 도출

- 주목할 점은 여타 사이버, 우주 국제법 및 규범 형성 과정이 국가·진영 간 대립으로 인해 상당한 시일이 소요되고 있으나, AI의 경우 상대적으로 빠르게 추진
 - AI 관련 최초의 구속력 있는 국제법인 ‘AI 프레임워크 협약’은 불과 1년여의 협상 기간을 거쳐 현재 EU 및 유럽의 다수 국가와 미국, 일본, 캐나다, 이스라엘이 서명을 완료해 곧 발효될 예정
 - 동 협약은 AI 시스템의 설계부터 폐기까지 전 생애주기 동안 인권, 민주주의, 법치주의 보호와 기술혁신 촉진을 목표
 - 협약은 투명성, 책임성, 평등 및 차별 금지, 개인정보 보호 등 여러 원칙을 제시하며, 이 원칙들이 AI 시스템 전 과정에서 준수되도록 요구
 - ※ 적용 대상은 인권, 민주주의, 법치주의 침해 가능성이 있는 공공 당국과 민간 행위자이나, 국가 안보 및 국방 관련 활동이나 미사용 AI 연구·개발 활동은 제외
 - 동 협약은 세계 최초의 법적 구속력을 갖춘 AI 협약으로서, 유럽평의회 비회원국 및 비국가 행위자들의 참여를 권장하며 글로벌 표준으로 자리매김할 가능성이 유력

→ AI 시스템의 특정 기술적 요소는 미국 등 개별 국가 차원의 국가 정책과 법제가 양자 및 소다(小多)자 형태의 규범 문서를 통해 확산

- 앞서 언급한 AI 시스템을 구성하는 데이터, 컴퓨팅, 알고리즘 이슈들은 각기 소다자 및 국제기구 등을 통해 국제 규범화 되고 있으며 글로벌 공급망 및 기술 표준화 관점에서 전개
 - 실제로 미국은 관련 분야의 수출통제, 기술이전, 투자감시 등 다양한 법안을 추진하고 있으며, 이를 양자, 소다자, 글로벌 차원으로 확산
- AI 군사적 이용 및 무기체계 관련 국제법 창설 여부는 주요 AI 강국들이 모두 소극적이어서 조기에 조약 창설 가능성은 쉽지 않으나, 연성법 형태의 국제규범 논의는 보다 구체화될 가능성

→ 신기술 분야에서 미·중 규범 경쟁이 심화됨에 따라 AI 영역에서도 국제 안보 맥락에서 미·중 간 규범 경쟁이 가시화

- 국제 안보 차원의 규범화는 위험, 위협, 취약성의 세 가지 요소를 고려하여 규범설계를 해야 하므로, 각 요소를 구체화하는 과정에서 국가·진영 간 대립이 심화될 전망
- AI 부문에서 미국의 규범 형성의 경우 바이든 행정부 후반부부터 미국의 입장 변화가 보이는데, 사이버, 우주 분야와 같이 진영 경쟁 차원에서 인식하기 시작
 - 이러한 미국의 입장 변화 배경에는 미국의 AI 패권적 지위에 대한 위기, AI 범용성에 따른 확산은 경쟁국의 비대칭전력 강화 등에 기인

- ‘반(反)글로벌주의’적인 트럼프 2기 행정부 초반에는 국제규범 창설에 소극적일 수 있으나, AI 글로벌 거버넌스에서 중국의 주도권을 견제하는 차원에서 임기 후반부에는 국제규범 논의를 재점화할 가능성 상존

〈 글로벌 AI 거버넌스 관련 미·중 간 입장 차 〉

주요 현안 및 쟁점	미국	중국
기본적 목적과 원칙의 방향성	• AI의 책임 있는 이용과 안전성, 투명성, 인권 강조	• AI의 지속개발과 역량 강화, 격차 해소 강조
범위	• 상업적 이용과 군사적 이용 구분 • 국제 안보 맥락에서 논의 가능	• 상업적 이용 중심의 논의 선호 • 국제 안보 차원의 논의 소극적
핵심 의제	• 데이터의 자유로운 이동, 컴퓨팅 기술 통제, 폐쇄형 AI 알고리즘/모델	• 데이터 주권, 컴퓨팅 글로벌 공급망 안전성, 개방형 AI 알고리즘/모델
국제규범 형성	• 조약 창설 반대 (군사 영역) • 연성 규범 및 정치적 규범 선호	• 조약 창설의 반대 (상업 및 군사 영역) • 자국 주도의 글로벌 이니셔티브 확산
국제기구 설립	• 정치적 형태의 회의체를 선호 • 미국과 유사 입장국 중심 운영	• 법률적 형태의 국제기구 설립 및 관련 국제기구 자국 유지 추진

출처 : 저자 작성

3 트럼프 제2기 행정부 AI 전략

➔ (개요 및 배경) 트럼프 행정부의 AI 전략의 기초는 미국의 글로벌 기술 패권 확보와 국내 AI 혁신을 목표로 한 국가 전략에 기반

- AI 기술은 '10년대 중반부터 급속히 발전하였으며, 이를 둘러싼 주요국 간 기술 경쟁도 심화됐으며, 특히 중국의 “차세대 인공지능 발전 계획(17)” 발표 이후, 미국 내에서는 AI 기술 패권 경쟁에 대한 위기의식이 확산
 - 트럼프 행정부는 '19년 2월, ‘미국 AI 이니셔티브’를 담은 행정명령(E.O. 13859)를 발표하며, 인공지능 분야에서 미국의 기술적 우위를 유지하기 위한 국가 차원의 정책적 대응을 본격화
 - 이 명령은 AI 연구개발 투자 확대, 인재 양성, 윤리적 기준 마련, 데이터 공유 및 접근성 향상 등을 골자로 하였으나, 그 구체적 실행은 제한적
 - 이후 트럼프 1기 임기 말 '20년에는 비(非)국방 분야 최초로 국가 AI 전략을 법제화한 「국가 AI 이니셔티브 법」을 제정하고, AI 연구개발, 협업, 인력 양성, 표준 및 정책 조율을 포괄하는 국가 차원의 제도 기반을 구축
- 트럼프 정부의 AI 정책은 철저히 민간 주도 혁신 모델에 기반하고 있으며, 이는 기술개발의 속도를 저해할 수 있는 과도한 규제를 배제하고, 산업계 중심의 시장 기반 경쟁을 촉진하려는데 초점

- AI가 미국 경제성장의 핵심축이 될 것으로 인식함에 따라, 연방정부는 정부 중심 연구개발보다 민간 기술의 확산과 응용에 집중
- 특히 국방 및 정보 분야에서는 AI를 차세대 전력 증강 수단으로 간주하면서 국방고등연구계획국(DARPA)은 'AI Next' 캠페인을 통해 방위산업에서의 응용 및 확대를 시도
- '25년 1월, 트럼프는 'AI 분야에서의 미국 리더십에 대한 장벽 제거'(E.O. 14179)에 서명해 바이든 정부 시기 도입된 AI 안전 프레임워크와 신뢰성 확보 규제 체계를 전면 철폐
- 동 행정명령은 'AI 행동계획(Action Plan)' 수립을 지시하며, 연방정부의 기술개발 규제를 완화하고, 민간 혁신을 저해하는 DEI(Diversity, Equity, Inclusion) 및 윤리 평가 프레임워크를 배제
- 트럼프 행정부의 AI 전략 배경은 AI를 단순한 기술로 보지 않고, 지정학적 패권과 산업 경쟁력의 중심에 놓인 '전략 자산'으로 인식
- 트럼프 대통령은 AI 행동계획과 관련 3개의 행정명령에 서명('25.7.23)
 - ※ AI 실행계획 발표와 동시에 데이터센터 허가 절차 완화(E.O. 14318), AI 이념 편향성 방지(E.O. 14319), 미국산 AI 기술 스팩 수출 촉진(E.O. 14320)의 3건의 행정명령에 서명하여 실행계획의 주요 권고사항 이행에 착수
- 상기 행정명령들은 행정부의 AI 행동계획 발표 직후에 발표
- AI 행동계획에서는 각각 세 가지 AI 정책 우선순위(△AI 인프라 구축, △미국 AI 기술의 글로벌 확산, △AI 모델에서 이념적 편향 제거)를 강조
- 트럼프 행정부의 'AI 행동계획'은 미국이 AI 분야의 리더십을 토대로 국가 경쟁력, 경제적 독립, 국가 안보의 원천으로 삼겠다는 강력한 의지를 표명한 것으로 평가
- 상기 행정계획과 행정명령은 핵심기술 분야에서 미국의 지위를 강화하기 위한 방안을 담은 방대한 양의 문서로 미국이 AI 분야에서 유일한 진정한 경쟁자인 중국에 대한 대응을 위해 설계
- 트럼프 행정부의 전략적 목표는 미국 AI 산업을 지원하기 위해 전체 자원을 효율적으로 활용하여 미국 AI 기술 패키지의 수출을 촉진하는데 초점
- ➔ (바이든 정부와의 차이) 트럼프 행정부의 AI 전략은 전임 바이든 행정부와 △AI를 보는 관점, △정책 추진 방식, △윤리 및 규제 접근, △국제협력 전략 등에서 상이한 관점을 바탕으로 전략을 수립
- 전략 및 정책 목표의 차이와 관련해서 바이든 정부는 AI를 사회적 신뢰 기반의 공공기술로 인식하고, 윤리와 공정성을 중심으로 한 포괄적 규제의 틀을 마련



※ '23년 발표된 'AI Bill of Rights'는 AI가 시민의 권리를 침해하지 않도록 보호하는 원칙을 제시하였으며, 미국 국립표준기술연구소(NIST)는 '신뢰할 수 있는 AI 프레임워크'를 개발해 연방 정부 조달과 평가 기준에 반영

- 트럼프 행정부는 AI를 경제성장 및 지정학적 우위를 위한 산업 전략 기술로 정의
 - 시장 주도와 규제 완화가 중심축으로 정부조달 기준에서 윤리·공정성 요건을 삭제하고, 기술적 효율성과 비용 대비 효과를 우선시
- 바이든 행정부는 DEI 요소를 강화하고, AI의 편향성과 차별성을 최소화하기 위한 윤리 가이드라인을 제도화
 - 연방 기관은 AI 도입 시 공공 참여, 영향 평가, 설명 가능성 확보 등을 요구
 - 반면, 트럼프 행정부는 이러한 접근을 '이념적 간섭'으로 간주하고, 윤리 지침 및 DEI 규정을 철폐하였으며, '이념적으로 중립적인 AI'를 새로운 표준으로 강조
- 국내 추진체계 및 거버넌스 구조의 경우, 바이든 정부는 AI 윤리위원회, RAI(Responsible AI) 센터 등 다층적 협의체를 운영하였으며, 연방 인재 채용 시에도 다양한 배경과 분야의 전문가를 포괄
 - 바이든 행정부 시기 백악관 과학기술정책실(OSTP)은 이러한 「국가 AI 이니셔티브 법」에 의거하여 '국가 AI 이니셔티브 실(National Artificial Intelligence Initiative Office, NAIIO)'을 출범시켜 국가 인공지능 정책을 마련
 - 트럼프 행정부는 이러한 구조를 축소하고, 실리콘밸리 및 기술기업 인사 중심의 자문위원회를 설치
- 전임 바이든 정부는 동맹 및 파트너 국가와의 협력을 강조하면서 다자적 접근을 중시한 반면, 트럼프 행정부는 일방주의적 노선 추구
 - 바이든 행정부는 EU, OECD, G7, UNESCO 등과 협력하여 글로벌 AI 규범 형성에 참여하였고, 민주주의 연합 내 AI 윤리 기준을 공유하려는 노력을 지속
 - 트럼프 행정부는 반대로 다자주의보다는 미국 단독 기술 리더십을 강조하며, '글로벌 규범'보다는 '미국 중심 블록'을 우선시

➔ **(전략의 주요 특성)** 트럼프 행정부의 AI 전략은 기본적으로 '미국 우선주의(America First)' 기술 정책의 연장선상에서, 친시장·반규제 중심의 기술혁신 모델을 표방

- 트럼프 행정부는 기술혁신의 주요 원동력이 민간 기업에 있다고 판단하고, 정부는 최소한의 개입을 통해 AI 개발 환경을 조성해야 한다고 강조
 - AI 관련 연방 규제안 마련은 지연되었고, 주(州) 정부 단위에서 추진되던 AI 책임성 및 윤리 기준과도 일정한 충돌이 발생

- 트럼프 1기 행정부는 AI를 안보 및 경제 전략의 핵심 기술로 규정
- 트럼프 행정부 AI 정책의 또 다른 특성은 ‘이념적 중립성(Unbiased AI)’이라는 새로운 정책 프레임을 제시
 - 이는 트럼프 진영에서 주장해 온 ‘Woke AI’ 담론과 연결되며, 알고리즘 및 모델 학습 데이터에서 DEI 요소를 제거하고, 이념적 편향성이 없도록 하자는 취지로 제시
 - '25년 발표된 행동계획에서는 연방정부가 사용하는 AI 모델에 대해 ‘진실되고 이념적으로 중립적’일 것을 요구하는 새로운 표준을 제시
 - 연방정부 내 DEI 및 기후 기반 지침을 철폐하면서 AI 인프라(데이터센터, 연산 인프라 등) 구축 시 NEPA(국가환경정책법) 검토 요건을 완화하여 에너지 소비가 높은 고성능 AI 개발을 장려
- 트럼프 행정부 AI 전략은 중국과의 AI 기술 패권 경쟁에 대응하기 위해 기술 안보와 국제 경쟁력 강화를 핵심 목표로 설정
 - 트럼프 행정부는 AI를 국가 안보와 직접 연결된 ‘국가전략기술’로 간주하였으며, 전임 바이든 행정부와는 달리 동맹국과의 협력을 확대하기보다는 미국 중심의 기술 블록 형성을 선호
 - 특히 AI 관련 기술(모델, API, GPU, LLM 등)의 해외 이전 제한과 수출입 통제 완화는 ‘전략적 우위 확보’라는 목적하에 병행 추진



〈 역대 미 행정부의 주요 전략·정책 (2013~) 〉

오바마 2기 (2013 ~ 2016)	
Preparing for the Future of AI, National AI R&D Strategic Plan (2016.10.12.)	공익을 위한 AI 활용, 연방정부 AI 활용, 규제혁신, AI 인재 양성, 고용 시장 영향 파악, 안전·거버넌스 설계 등 파악, 안전·거버넌스 설계 등 AI의 사회적 영향 정책 방향 제시
Artificial Intelligence, Automation, and the Economy (2016.12.20.)	AI 및 자동화가 노동시장에 미칠 충격과 경제적 과제 진단 및 대응을 위한 정책 권고
트럼프 1기 행정부 (2017 ~ 2020)	
E.O.13859 Maintaining American Leadership in AI (2019.2.11.)	AI 연구개발 투자 확대, 연방 데이터컴퓨팅 자원 개발, 기술 표준 정립, 인력 양성, 국제 협력 등 5대 중점 방향 설정
National AI Initiative Act of 2020 (2020.12.27.)	AI 연구개발, 협업, 인력 양성, 표준 및 정책 조율 등 국가 차원의 제도 기반 구축
바이든 행정부(2021 ~ 2024)	
CHIPS and Science Act (2022.8.9.)	핵심 기술 인프라에 대한 대규모 연방 투자 프로그램 출범, 중국과의 기술 패권 경쟁 대응을 위한 산업 기반 마련
Blueprint for an AI Bill of Rights (2022.10.)	안전성, 차별 방지, 프라이버시 보호, 인간 개입 보장 등 5대 원칙 제시하는 등 민간 영역의 책임 있는 AI 개발 촉구
E.O.14110 Safe, Secure and Trustworthy Development and Use of Artificial Intelligence (2023.10.30.)	연방정부 차원의 포괄적 AI 감독 체계 구축
트럼프 2기 행정부 (2024 ~)	
E.O.14148 Initial Rescission of Harmful (2025.1.20.)	바이든 행정부 AI 규제 정책 철회
E.O.14179 Removing barriers to American Leadership in AI (2024.1.23.)	인류의 지속 가능한 번영, 경제적 경쟁력 강화, 국가 안보 증진을 통해 미국의 글로벌 AI 패권 지속·강화.
AI Action Plan (2023.7.23.)	① AI 혁신 가속화, ② AI 인프라 구축, ③ 국제 AI 외교 및 안보 선도 등 3대 전략 축, 30대 실행 과제, 103개의 정책 권고사항으로 구성됨
E.O.14318 Accelerating Federal permitting of Data Center Infrastructure (2025.7.23.)	AI 데이터센터 및 핵심 인프라의 신속한 구축을 위해 규제 완화, 연방 토지 제공, 재정 지원을 골자로 한 대규모 산업 계획 수립
E.O.14319 Preventing Woke AI in the Federal Government (2025.7.23.)	연방정부가 조달하는 AI 언어모델에서 다양성, 형평성, 포용성 등 이념적 편향을 배제하고, 진실성과 중립성을 보장하는 AI만 사용 토록 규정
E.O.14320 Promoting The Export of the American AI Technology Stack (2025.7.23.)	미국산 AI 기술 스택의 수출을 촉진하고, 동맹국과의 기술 협력 및 지정학적 우위 확보를 위한 '미국 AI 수출 프로그램(American AI Exports Program)'을 신설하여 연방정부 차원의 정책·재정·외교 자원을 총동원

출처 : European Parliament

4 향후 전망 및 시사점

→ 미·중 등 주요 AI 강국들은 글로벌 차원의 AI 거버넌스 논의에 소극적이었으나 향후 국제 안보 맥락에서 글로벌 AI 거버넌스 경쟁이 본격적으로 전개

- 미·중 간 AI 패권 경쟁은 주로 국내 차원의 AI 혁신전략을 통해 구체화되었으나, 미·중 간 글로벌 거버넌스 경쟁으로 전이될 것으로 전망
 - 글로벌 이슈에 소극적인 트럼프 행정부의 특성상 단기적으로는 거버넌스 논의의 휴지기가 도래할 수 있으며 상업적 이익 강조 및 국가주의적 접근이 확대
 - 기존 AI 거버넌스 논의는 개발, 안전성, 국제 안보, 군사적 이용, 무기체계 관점에서 분절적·파편적으로 논의되었으나 추후 국제 안보 차원으로 수렴 가능
 - 분절적인 거버넌스 체계의 연결 가능성이 주요한 현안으로 부상될 것으로 전망
- 미·중 AI 거버넌스 경쟁은 글로벌 사이버 거버넌스 경쟁 구조와 유사성을 갖고 있지만, 미·중 전략 경쟁 차원에서는 거버넌스 구축 시 공통의 이해 상존
 - AI는 단순한 기술이 아니라 글로벌 패권 경쟁의 수단으로 부상하고 있으며, 특히 미·중 간 전략적 경쟁이 글로벌 AI 거버넌스 구축에 있어 다자적 협력에 장애 요인으로 작용
 - AI의 파괴적(disruptive)인 특성으로 인해 역으로 국제 안보 관점의 거버넌스 구축 시 미·중 간 협력을 도모할 가능성 상존
 - 과거 미국·소련 간 우주·핵 경쟁이 격화되면서도 다자적 우주, 핵 레짐 구축 시 협력했던 사례와 같이 실제, 미·중 간에는 AI 전략 대화를 통해 AI 거버넌스 관련 공통 이해를 제고
- 국제 안보 중심적 논의는 협력보다 경쟁 중심의 글로벌 거버넌스 양상이 전개될 가능성이 우세
 - 공동 규범보다는 자국 우선 전략이 강조되면서 중립적 국제 거버넌스 기구 설립 난항이 예상
 - 거버넌스 형성의 도전 요소인 △국제협력 및 조율 부족, △국가 간 경쟁 및 신뢰 부족, △거버넌스 체계의 분열, △AI 거대 기업으로 영향력 집중, △AI 위협에 대한 교육 및 대중의 이해 부족 등의 한계 지속될 전망
 - 거버넌스 주요 의제의 경우, ‘인간 중심 접근(human centered approach)’과 ‘위험 기반 접근(risk based approach)’의 기존 논의에서, 향후 국제 안보에 위협이 되는 요소를 식별하고 이에 대응하는 신뢰 구축 조치와 규범 형성이 주요 현안으로 대두될 것으로 예상



- ➔ AI는 비대칭전력의 핵심 수단이자 모든 형태의 무기체계에 접목될 수 있어 군사적 이용이 고도화될 것이며 이에 따른 군비경쟁 역시 심화될 전망
- AI 기술 선진국들은 AI 군사력 증강을 적극적으로 추진하고 있으며 기술 수준에 따라 군사적 이용 범위 역시 증대
 - 글로벌 AI 거버넌스 맥락에서는 군사적 이용 시 안전성 확보를 위한 거버넌스 및 규범 형성 논의가 빠르게 진전될 것으로 전망
 - 군사적 이용에 대한 규범 논의는 기존 LAWS 거버넌스의 중복성을 피하기 위해 논의의 지향점(연성 규범, 조약 창설 등)을 달리하여 병행될 것으로 예상
- AI의 군사적 이용의 경우 여타 신기술의 군사적 이용과 차이가 있으며 거버넌스의 방향성도 사이버·우주와는 상이
 - 구체적인 세부 의제는 △군사 분야 AI 위험 요소에 대한 공동의 평가, △최적 관행 공유, △공통 용어집 개발, △지정학적 긴장과 AI 거버넌스 논의를 위한 국제협력 등이 쟁점
 - ※ AI의 군사적 이용의 경우 공격-방어의 불균형이 상당하므로, AI의 군사적 개발에 있어 90% 정도가 방어 능력보다는 공격 능력 배양에 치중
- ➔ 최근 발생한 미토스(Mythos) 쇼크와 같이 AI 기반 사이버 안보의 새로운 위협 요인으로 부상하고 있으며 기술 및 수출통제 관점에서 접근
- 미토스 쇼크는 AI 자체의 안전성, 편향성, 인권 차원의 위험과는 맥락이 다른 것으로 기존 사이버 위협의 고도화·복잡화를 심화시키는 사이버 안보 영역으로 진화·전이
 - 사이버 안보는 신기술의 발전에 따라 새로운 위협 요인으로 주요 기반 시설에 대한 취약성이 지속적으로 노출되어왔으며 이번 미토스 사태의 경우 사이버와 AI 기술이 연계된 새로운 형태의 위협 요인이 부각
 - 향후, 이에 대한 안전성 평가 및 대응 조치 역시 복잡하고 고도화될 것이 확실
- 미토스 쇼크에서 시사하는 것은 경제·기술 안보적 차원에서 취해진 조치들은 향후 한국을 비롯한 동맹·파트너 국가에 새로운 도전 요인으로 작용
 - 글로벌 AI 경쟁 맥락에서, 경쟁국에 대한 통제와 이에 대한 공조에 대한 압박을 받을 것은 물론 누구나 미국의 전략 자산에 대한 통제 대상이 될 가능성이 상당
 - 기존 첨단 반도체 및 하드웨어는 물론 소버린 AI 전략 차원에서 추진 중인 첨단, 고성능 AI 모델 개발·활용·수출입 등 AI 혁신 생태계 구축 전반에 걸쳐 장애 요인으로 기능

- 소버린 AI 전략·정책은 기술 주권을 추구하려는 목표가 있기 때문에, 한국의 국가 AI 전략에 있어 혁신과 규제라는 이분법적인 접근보다는 기술 표준화 맥락에서 혁신의 방향이 규제 기준과 양립할 수 있는 방안을 모색
- 미·중 간 AI 기술 패권 경쟁이 심화됨에 따라 중국에 대한 경쟁 우위를 확보하기 위한 수단으로 수출통제 거버넌스를 양자, 소다자, 글로벌 차원으로 확대·강화할 것으로 전망
- 최근 미국 관련 기업의 이해를 반영하여 각국을 세 그룹(동맹국 및 파트너 국가, 중국과 러시아 등 적대국 및 비우호 국가, 기타 일반 국가)으로 분류하여 AI 반도체 수출 거래 방식을 차별적으로 규제하려는 정책 지속 가능
- 러시아의 우크라이나 침공, 미·중 전략 경쟁 심화에 따라 국제 비확산체제는 중요한 전환점에 있으며, AI 기술 통제를 위한 포스트 바세나르(Post-Wassenaar Arrangement)* 체제 출현 가능성 상존
- * 기존의 다자간 수출통제 기구인 바세나르 체제(WA)의 한계를 극복하고, 급변하는 첨단 기술 환경과 글로벌 안보 위기 속에서 새로운 다자간 첨단기술 수출통제 연합을 구축하려는 움직임 또는 그 구상