THE
AI
REPORT

미국 AI 국가안보각서 (AI NSM) 분석 및 시사점

2024

한국지능정보사회진흥원

「The AI Report」는 인공지능 기술·산업·정책의 글로벌 이슈와 동향, 시사점을 적시에 분석, 인공지능 현안에 빠르게 대응하고 관련 정책을 지원하기 위해 한국지능정보사회진흥원(NIA)에서 기획·발간하고 있습니다.

▶ 발행인 : 황 종 성

▶ 작 성
- 한국지능정보사회진흥원 인공지능정책본부 AI정책연구팀 김태순 책임연구원(ts_kim@nia.or.kr)
- 한국지능정보사회진흥원 인공지능정책본부 AI정책연구팀 이상은 주임연구원(slee@nia.or.kr)

미국 AI 국가안보각서(AI NSM) 분석 및 시사점

NIA AI정책연구팀 김태순 책임연구원, 이상은 주임연구원

■ 본 보고서는 美 바이든 정부가 AI를 국가안보에 책임있게 사용하는 데 필요한 지침과 이행사항을 담아 발표한 'AI 국가안보각서(AI NSM, National Security Memorandum on AI)'*를 요약·분석

* Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence(현지 시간 기준 10월 24일(목), 백악관 홈페이지에 문서 게재)

AI 국가안보각서 수립 배경 및 목적

○ 제이크 설리번 美 국가안보보좌관은 AI 국가안보각서(이하 'AI NSM') 발표 브리핑에서 이하의 수립 배경, 목적을 설명

- **(AI 행정명령의 이행)** 'AI 행정명령('23.10) 4.8항(AI에 관한 국가보안각서 개발)' 이행을 위해 부통령, 국무·재무·국방장관, 법무부장관 등 26개 연방정부 기관장이 서명
- **(AI의 특징)** AI가 다른 기술 도약(전기화, 핵무기, 우주 비행, 인터넷)과 다른 세 가지 핵심적인 측면으로 **▲빠른 발전 속도, ▲AI의 성장 궤적에 대한 불확실성, ▲정부가 아닌 민간 기업이 AI 개발을 주도**하고 있다는 특징을 지적

  ※ 정부는 모든 시나리오에 대비하고 미국의 혁신 생태계를 보호할 국가 안보 정책을 구축해야 하며, 특히 민간 부문이 주도하는 현재 AI 생태계에서 정부는 기술 관리자·배포자로서 역학 관계의 의미를 명확히 파악해야 한다고 강조
- **(AI와 국가안보)** AI가 계속해서 더욱 강력하고 범용적으로 발전할 것으로 예상되는 상황에서 "*향후 몇 년 동안은 AI보다 더 중요한 기술은 없을 것*"이라고 강조하며, **AI와 국가 안보 관점에서 미국의 현주소를 진단**

  ※ AI는 국가 안보와 관련한 거의 모든 주요 영역(핵물리학, 로켓 공학, 스텔스 기술, 핵무기, 언론 자유 등)에 영향을 미칠 가능성 보유

  · 美 바이든 대통령의 AI 행정명령은 전 세계에서 AI에 대한 가장 포괄적인 조치이며, 이를 통해 AI 인재·하드웨어·인프라·거버넌스를 강화하기 위한 노력 지속하는 점은 긍정적

  · 하지만 이것으로 미국의 우위가 보장되는 것은 아니며 현재의 발전 속도를 유지하는 것을 넘어서 미국의 경쟁자들보다 더 빠르게 국가 안보 강화를 위해 신속하고 폭넓게 AI 배치할 필요

  ※ 현재 미국은 최고의 AI 모델을 보유하고 있지만, 경쟁사가 더 빠르게 배포한다면 미국 국민, 군대와 미국의 동맹국, 파트너를 상대로 AI 역량을 사용하는 데 우위를 점하게 될 수 있다고 우려 표명
- **(전략적 기습)** 미국 기업들이 전세계적으로 AI 분야를 주도하고 있는 상황에서 중국을 포함한 **'경쟁국들로부터의 전략적 기습**(risk of a strategic surprise by our rivals)'의 위험이라는 도전과제 해결을 위해 AI NSM를 마련했다고 설명

○ 특히, 미국-중국의 본격화되는 AI 경쟁은 AI NSM 수립의 핵심적 배경

- **(AI 반도체)** 미국은 대중 반도체 장비 및 기술 수출통제 조치를 발효하는 등 규제를 점진적으로 확대·강화
  · ('22.10) 美 상무부는 특정 반도체(첨단 컴퓨팅 칩, 슈퍼 컴퓨터 부품), 장비, 관련 소프트웨어 및 기술 등에 대한 수출 통제 조치 발표
  ※ 엔비디아, AMD에 중국이 군사용으로 사용할 우려가 있는 AI 반도체 수출 금지
  · ('23.5) 반도체법(Chips and Science Act) 가드레일 조항 세부 규정안을 통해 대중국 반도체 통제 강화
  ※ 보조금을 받은 기업에 대해 중국 등 우려 국가에서 허용치 이상으로 반도체 생산능력을 확장하는 경우 보조금 전액 반환 규정 등
  · ('23.10) 수출통제 조치를 강화하는 수출관리규정(EAR) 개정안 발표, ▲AI 칩 및 장비 통제 대상 확대, ▲우회 수출 적용 대상 강화, ▲우려거래자 목록 추가 등 대중 첨단 반도체 제재 본격화
  ※ 엔비디아에 대중국 첨단 AI 반도체 수출 통제, 2023년 미국산 기술이 포함된 반도체 제조 장비 수출 금지, AI와 슈퍼컴퓨터에 사용하는 반도체 칩 수출 제한 등
  · ('24.6) GAA(Gate All Around)·HBM(High Bandwidth Memory) 등 중국의 첨단 AI 반도체 기술 접근을 제한하는 방안을 검토하며, 일본, 네덜란드 등 동맹국과의 협력 강화 추진 및 공동 대응 촉구
  · ('24.10) 美 재무부는 6월부터 검토 및 의견수렴 과정을 거쳐 반도체·마이크로전자기술·양자컴퓨팅·AI 등 분야에서 우려 대상국인 중국(홍콩, 마카오 포함)에 대한 투자를 제한하는 최종 규칙 확정·발표('24.10.28)
  ※ 「우려 국가의 특정 국가 안보·기술 제품에 대한 미국 투자 제한 행정명령(EO 14105)」('23.8.9)에 따른 것으로, '25.1.2부터 시행
  ※ *"이 최종 규칙은 미국의 투자가 미국의 국가 안보를 위협하는 자들의 핵심 기술 개발을 돕는 것을 막기위한 조치"* (폴 로젠 美 재무부 투자 안보 담당 차관보, '24.10.28)

- **(미·중 논의)** 미국, 중국은 AI 기술의 군사적 사용, 안전성, 위험 등에 대한 논의에서 전략적 입장을 표명해오고 있음
  · ('23.11) 美 샌프란시스코 비공개 회담에서 바이든 대통령, 시진핑 주석은 AI 위험과 안전에 관한 대화에 합의
  ※ AI 기술이 실존적 위협이 되지 않도록 하자는 의견 등을 교환
  · ('23.11) 美 국무부는 「AI와 자율성의 책임있는 군사적 이용에 관한 정치적 선언」 발표
  ※ 한국, 미국 등 60여 개 국가가 참여하고 북한·중국·러시아는 선언에 미참여
  · ('24.2) 중국, 러시아는 군사적 AI 사용에 대한 협력 강화 약속
  · ('24.3) 미국은 만장일치로 통과된 최초의 UN 총회 AI 결의안을 후원하였으며, 중국을 공동 후원국으로 포함
  · ('24.5) 스위스 제네바에서 미국, 중국 정부 관계자들은 첫 AI 관련 고위급 회담 개최
  ※ 미 국무부는 AI의 군사적 이용에 관한 정치적 선언에 대한 중국·러시아의 참여를 촉구하는 한편, 중국은 AI 분야에서 미국의 대중국 탄압에 대해 단호한 입장을 밝히며 글로벌 거버넌스에 대한 주요 채널로서 UN의 역할 옹호(AI타임즈, '24.5)

☞ 제이크 설리번은 브리핑에서 *"AI 위험을 더 잘 이해하기 위해 중국 및 다른 국가들도 기술에 대한 대화에 참여해야 한다"*고 언급하면서도, *"이러한 회의가 중국이 AI를 사용하여 국민을 억압하고, 잘못된 정보를 퍼뜨리고, 미국과 동맹국, 파트너국의 안보를 훼손하는 방식에 대한 우리의 깊은 우려를 줄이는 것은 아니다."*라고 강조

AI 국가안보각서의 목표 및 주요 내용

○ AI NSM은 ▲안전하고 신뢰할 수 있는 AI 개발의 글로벌 선도(미국의 리더십 유지), ▲국가안보를 위한 AI 활용(AI 시스템 도입 가속화) 및 민주적 가치 보호, ▲책임있는 국제 AI 거버넌스 구축을 목표로 설정

※ 제이크 설리번의 브리핑(백악관, '24.10.24) 때 발표된 내용을 기반으로 주요 내용을 아래에 요약·정리

- 미국의 AI 리더십 확보

· (인재) 비자 절차 간소화 등을 통한 신기술 분야 인재 유치를 위한 조치 확대

· (하드웨어·전력) 모든 국가 안보 기관에 칩 공급망이 안전하고 외국의 간섭으로부터 자유로울 수 있도록 지시, 기후 목표에 부합하는 방식으로 AI 데이터 센터 지원할 수 있는 청정 에너지 발전 시설 설계·허가·건설 지원

· (혁신을 위한 자금 지원) 초당적 지지를 기반으로 AI R&D 예산 확보 노력

· (기술 보호) AI 부문에 대한 첩보, 스파이 등 적대적 위협을 우선순위로 설정하고, 해당 위협에 더 많은 자원과 인력 투입

※ 민간 부문 AI 개발자와 긴밀한 협력을 통해 민간 부문의 기술을 안전하게 보호하기 위해 사이버 보안 및 방첩 정보를 적시에 제공하도록 도모

- 국가안보를 위한 AI 활용

· (협력) 각 기관의 선도적인 AI 기업 및 클라우드 컴퓨팅 제공업체와 같은 비전통적인 공급업체와 협력 강화

· (공유 컴퓨팅 리소스) 핵·바이오·사이버 보안까지 광범위한 위협에 책임있게 대처하면서 AI 도입을 가속화하고, 비용을 절감하며, 상호 학습할 수 있도록 공유 컴퓨팅 리소스 사용 장려

※ 안전하고 보안을 유지하면서 '책임 있게(responsibly)' AI를 개발·배포하는 것이 미국 전략의 중추

· (프레임워크) **국가 안보 분야에서 AI 위험 관리 약속에 관한 최초의 정부 차원 프레임워크 구축**(유해 편견과 차별 방지, 책임성 극대화, 효과적이고 적절한 인적 감독 보장 등)

※ 프레임워크는 ▲금지되고 영향력이 큰 AI 사용 사례 및 연방 직원에게 영향을 미치는 AI 사용 사례 식별, ▲영향력이 큰 것으로 식별된 AI 범주에 대해 충분히 강력한 최소 위험 관리 방법 수립, ▲영향력이 큰 AI 사용 목록화 및 모니터링, ▲효과적인 교육 및 책임 메커니즘 보장에 대한 네 가지 축으로 구성

· (책임성) 핵심 가치 준수, 차별 방지, 책임성 강화, 인간의 감독 보장 등의 원칙 확립

- 책임있는 국제 AI 거버넌스 구축

· 전 세계 사람들이 AI로 인한 혜택을 누리고 위험을 완화하도록 하기 위해서는 국제 규범과 파트너십 구축이 필요

※ 미국은 AI의 군사적 사용에 관한 정치 선언 발표('23.11), G7 주도의 최초 AI 국제 행동강령 개발('23.12), 블레츨리·서울 정상회의에서 20여 국가와 함께 AI의 원칙 마련('23.11, '24.5), 최초의 UN AI 결의안 채택('24.3) 등에 참여 및 주도

· 최첨단 AI 기술을 보호하는 한편, 전 세계의 AI 채택을 촉진하는 균형을 유지할 필요

※ 브리핑에서 **AI NSM과 별개로 새로운 글로벌 AI 확산 접근법에 대한 발표를 예고**: ▲최첨단 칩의 수출 관리 방안, ▲광범위한 AI 컴퓨팅에 대한 접근 보장 방안, ▲미국의 선도 AI 기업과 AI 혁명에 참여하고자 하는 전 세계 국가 간 파트너십 촉진 방안, ▲위험 방지 및 기회 창출을 위한 안전·보안 기준의 설정 방향 등을 포함

☑ 주요 대상

○ 국가안전보장회의(NSC) 정보 및 국방정책조정관 마허 비타르(Maher Bitar)는 *"AI NSM이 많은 독자들에게 공개되는 동시에, 일부는 기밀로 유지될 것"*이라고 언급('24.6)

○ 美 전략국제문제연구소(CSIS)는 AI NSM의 "독자"들 중 주요 대상을 아래 네 그룹으로 분석

- **(미연방 기관과 직원들)** AI NSM은 프론티어 AI에 대한 미국의 국가안보정책을 제시하고, 정책 실행을 위해 여러 연방기관에 정책적 명확성과 구체적인 임무를 부여
- **(미국 AI 기업)** 미국 국가안보 이익 증진에 있어 공공부문과 민간부문의 적절한 역할에 대한 바이든 행정부의 시각을 명확화하며, 정부가 민간부문의 AI 리더십을 어떻게 지원할 것인지, 국가 안보를 위해 민간부문에 무엇을 기대하고 요구하는지 포함

  ※ 설리번 보좌관은 브리핑에서 *"AI 개발을 주도하는 것은 정부가 아닌 민간 기업들"이*라고 언급(백악관, '24.10.24)

- **(미국의 동맹국)** 미국은 AI가 국가안보에 중요 요소로 대두됨에 따라 국가 안보를 위한 정책 조치를 단행

  · AI NSM을 통해 미국의 동맹국들은 미국이 왜 프론티어 AI 시스템에서 리더십을 확보하는 것을 중요하게 여기는지, AI 리더십 유지를 위해 특별한 조치들을 단행하는지에 대한 기준점을 제공

  ※ 美 행정부는 첨단 AI 시스템을 가능하게 하는 반도체 기술의 대중국 수출에 대한 포괄적인 통제 정책 발표('22.10)

  · 또한 AI NSM는 현재 미국이 동맹국 및 파트너와 맺고 있는 AI 생태계 내 협력적인 행보와 관련하여 미국의 동맹국과 파트너들이 미국 주도 AI 생태계에서 어떠한 역할을 할 수 있을지 제시

  ※ 美-UAE AI 관련 데이터센터 및 에너지 인프라 구축 협약 체결('24.9.23), Microsoft-G42 AI 개발 협약 체결('24.4.15)

  ※ AI NSM 4.1(h) 조항은 '미국의 동맹국과 파트너 네트워크는 경쟁자들에 비해 상당한 이점을 제공한다. '22년 국가안보전략(National Security Strategy)이나 후속 전략과 일관되게, 미국 정부는 선별된 동맹국 및 파트너들과 AI 역량의 공동개발과 공동배치를 적극적으로 가능케 하고 투자해야 한다'고 명시

- **(미국의 적대국과 전략적 경쟁자)** AI NSM는 미국이 글로벌 AI 리더십에서 가장 강력한 경쟁자인 중국과의 경쟁에서 우위를 점하기 위한 방안들을 설명하는 한편, 일부 조항들은 그간 미·중 간의 외교적 접근과 향후 잠재적 외교적 접근을 보완하는 것으로 해석

  ※ AI NSM과 함께 발표된 '국가안보의 AI 거버넌스와 위험관리 발전을 위한 프레임워크'는 '대통령의 핵무기 사용 개시 또는 종료 결정을 알리고 실행하는 데 중요한 행동에서 인간을 배제하는 AI 사용을 금지'하였는데, 이 내용은 미·중 AI 안전 회의('24.5)에서 제기되었을 가능성이 높은 것으로 분석

| 참고 | AI와 국가 안보에 대한 OpenAI의 접근

■ **OpenAI는 AI NSM이 발표된 당일('24.10.24), AI와 국가 안보에 대한 OpenAI의 접근법을 발표[1)]**

- **(배경)** OpenAI는 백악관의 AI NSM이 AI 분야에서 미국의 리더십을 유지하는 데 필수적이라는 데 동의하며, 민주주의 가치를 옹호하는 방식으로 AI를 발전시키려는 노력에 있어서 중요한 진전으로 평가
- **(민주주의 비전)** AI 잠재력을 활용하고 혜택을 공유하기 위해서는 민주적인 비전(democratic vision)이 필수적이며, 따라서 민주주의 국가들이 자유·공정성·인권 존중과 같은 가치에 따라 AI 개발 주도 필요
- **(국가안보활용사례)** OpenAI는 위 비전에 부합하는 자사 AI의 국가 안보 분야 활용 사례 제시
  - 美 고등연구계획국(DARPA)이 개최하는 'AI 사이버 챌린지(AIxCC)'에서 앤트로픽, 구글, MS, OpenAI는 DARPA와 협력하여 챌린지 참가자들이 최첨단 사이버 보안 시스템을 개발할 수 있도록 지원('23.8)[2)]
  - 美 국제개발처(USAID)는 OpenAI와 협력하여 직원들의 행정 부담을 줄이기 위해 연방기관 최초로 ChatGPT 사용
  - 로스 알라모스 국립연구소와의 생명과학 연구 파트너십을 기반으로 미국 국립연구소와의 협력 강화

■ **OpenAI의 정책과 가치**

- **(민주적 가치)** AI가 자유 증진, 개인의 권리 보호, 혁신을 촉진하는 방식으로 개발되고 사용
  - 이를 위해 기술에 대한 접근을 민주화하고 경제적, 교육적, 사회적 혜택은 극대화하기 위한 실질적 조치 취해야 할 것
- **(안전)** 위험 완화, 보안 강화, 인권 보호에 AI가 사용되기를 희망
  - 모든 잠재적 애플리케이션을 엄격히 평가하여 원칙에 부합하는지 확인
- **(책임감(Responsibility))** AI는 공익을 위해 사용
  - 기본권을 침해하는 데 사용되는 AI를 금지하고, 특히 국가안보와 같은 민감 영역에서 모든 잠재적 파트너십에 이를 엄격히 적용
- **(책임성(Accountability))** AI 시스템은 책임성을 핵심으로 개발 및 배포되어야 함
  - 특히 정부 및 국가 안보와 관련된 모든 AI 애플리케이션은 감독, 명확한 사용 지침, 윤리 기준의 적용을 받아야 함

■ **향후 전망**

- AI NSM 발표로 미국과 동맹국 간 국가안보 분야에 대한 다양한 활동이 전개될 전망
  - 과학 연구, 물류 강화개선, 번역 및 요약 작업 간소화, 시민 피해 연구 및 완화에 기술 적용 가능
  - 위 분야에 대해 수행하는 모든 작업에 대해서는 엄격한 내부 검토 절차를 거칠 것

1) https://openai.com/global-affairs/openais-approach-to-ai-and-national-security/
2) https://www.darpa.mil/news-events/2023-08-09

구성

○ 총 8개 섹션으로 구성되었으며 1~2 섹션은 정책 배경 및 목적, 3~5 섹션은 주요 이행 내용, 6 섹션은 실행과 보고체계, 7~8 섹션은 정의 및 일반 조항을 다룸

※ 비공개 기밀 부록 포함

[표 1] 미 'AI 국가 안보 각서' 주요 이행 내용(3~5 섹션) 요약

주요 내용		
3.	**미국의 기초 AI 역량 증진 및 확보** ▶ 첨단 AI 시스템 개발에서 미국의 리더십 유지	
	3.1 미국 AI 개발의 진보, 혁신 및 경쟁 촉진	■ 모든 법적 권한을 활용해 AI·반도체 인재 유치 ■ 민감기술 분야 외국 인재 비자 행정처리 간소화 ■ AI 인프라 구축 승인 절차 간소화 ■ 인프라 관련 공공·민간 투자 장려
	3.2 외국의 정보 위협으로부터 미국 AI 보호	■ AI 공급망 파괴 경로 파악 및 위험 경감 조치 이행 ■ 외국 행위자의 지적 재산 위협 보호
	3.3 AI 안전, 보안, 신뢰성에 대한 위험 관리	■ 프론티어 AI 모델 사전 및 사후 안전 테스트 촉진 ■ AI 모델의 사이버·핵·방사능 부문별 안전성 평가 역량 개발 및 평가 진행
4.	**국가 안보 목표 달성을 위한 책임감 있는 AI 활용** ▶ 미국 국가안보 기관 전반에 걸친 프론티어 AI 시스템 도입 가속화	
	4.1 효과적이고 책임감 있는 AI 사용 활성화	■ (인재확보) 책임있는 AI 가속화를 위한 인재 채용·유지 정책 및 전략 수정 ■ (교육강화) 관리예산처(OMB)와 협의하여 AI 역량 제고를 위한 교육 및 훈련 마련 ■ (조달 시스템 개선) AI 조달·활용 실무그룹 구성을 통한 효과적인 조달 시스템 운영 ■ (정책개선) 국제법과 국내법 등에 따라 AI 관련 법적·정책적 프레임워크 검토 및 개선 ■ (동맹국 협력) 주요 동맹국과의 AI 및 AI 지원 자산의 공동 개발 및 공동 사용의 발전, 증대, 촉진의 타당성 평가
	4.2. AI 거버넌스 및 위험 관리 강화	■ AI 거버넌스·위험관리 프레임워크 수립 및 NSC 승인 ■ 각 기관 AI 최고책임자 지정 및 AI 거버넌스위원회 구성 ■ 국가안보시스템에서 AI 사용에 관한 구체적 지침 제공
5.	**안정적이고 책임감 있으며 전 세계적으로 유익한 국제 AI 거버넌스 환경** ▶ 미국 국가안보를 지원할 강력한 거버넌스 프레임워크 개발	
	■ 미국 국내 정책 및 국제 거버넌스와 일관된 국제 AI 거버넌스 규범 발전 전략 수립 - 유엔, G7 등 국제기구, 동맹국과의 관계 및 경쟁국 관계에 대한 지침 포함	

세부 내용(원문 요약)

1 정책

○ 본 각서는 행정명령 14110호(AI의 안전하고, 신뢰할 수 있는 개발 및 사용, '23.10.30)의 4.8항(AI에 관한 국가안보각서 개발)에 명시된 지침을 이행

- 미국 정부 국가안보시스템(NSS)에서 AI 모델과 AI 지원 기술을 적절히 활용하고 인권, 시민권, 시민자유, 프라이버시, 안전을 보호하는 지침 제공

※ 기밀 부록은 미국 국가안보에 위험을 초래하는 적대적 AI 사용 대응 등 추가 민감 국가안보 사안을 포함

○ 미국은 그간 잠수함, 항공기, 우주시스템, 사이버 도구 등 기술 전환기 때마다 첨단 기술을 이용해 우위를 점하려는 적의 시도를 추적하고 대응하기 위한 새로운 시스템을 개발 해옴

○ AI는 시대를 정의하는 기술로서 부상하고 국가안보 관련과 관계가 증가하고 있어 AI의 책임있는 활용에서 미국의 글로벌 선도가 필요함

※ 오용은 전 세계적으로 민주적 제도와 절차를 악화시키며 국제 질서를 약화 시킬수도 있음

○ AI 혁신 및 패러다임 변화가 민간 부문 주도로 이뤄지고 있으며, 특히 대규모 언어 모델에서 두드러지고 있어 정부는 이러한 AI 발전이 국가안보에 미칠 영향을 시급히 검토해야 함

○ AI 발전의 핵심 동인(알고리즘 개선, 컴퓨팅 성능 향상, 업계 투자, 데이터 확장)이 지속되어 더욱 강력하고 범용적인 AI가 예상됨에 따라 새로운 리소스와 인프라 체계가 필요

○ 정부가 산업계·학계와 협력해 AI 역량을 적극 활용하지 않으면 기술 우위를 상실하게 되며, 이는 국가안보와 민주적 가치 수호에 큰 위험이 될 것

○ AI 국가안보 리더십 확립을 위해서는 정부 조직 전반의 변화가 필요하며, 단일 기관이 아닌 범정부 차원의 통합적 접근과 인프라 재설계가 요구

○ 미국은 AI 발전 과정에서 인권과 민주적 가치를 보호하고, 국제 AI 거버넌스의 규범과 제도를 주도하는 글로벌 리더십을 발휘해야 함

○ 본 각서는 AI 생태계, AI 시스템의 안전성 확보, 국가안보를 위한 AI 활용 촉진, AI 오용 방지 등 미국의 AI 국가안보 정책 전반의 변화를 추진

2 AI 국가안보각서의 목표

○ ▲안전하고 신뢰할 수 있는 AI 개발의 글로벌 선도(미국의 리더십 유지), ▲국가안보를 위한 AI 활용(AI 시스템 도입 가속화) 및 민주적 가치 보호, ▲책임있는 국제 AI 거버넌스 구축을 목표로 설정

- 정부는 산업계, 시민사회, 학계와 협력해 AI 개발의 기초 역량을 확보하고, 글로벌 인재와 최고의 컴퓨팅 시설을 유치하며, AI 시스템의 위험을 평가하고 완화해야 함
- 정부는 국가안보 목표를 달성하기 위해 적절한 안전장치와 함께 강력한 AI(powerful AI)를 활용하며 혜택을 추구하면서도 그 한계를 인지해야 하며 모든 AI 활용 과정에서 투명성과 인권 등 민주적 가치를 보장해야함
- 정부는 안전하고, 보안이 확보되며, 신뢰할 수 있는 AI 개발과 사용을 촉진하고, AI 위험을 관리하며, 민주적 가치를 실현하고, 인권, 시민권, 시민의 자유, 프라이버시를 존중하며, AI의 전 세계적 혜택을 증진하는 국제 AI 거버넌스를 발전시키기 위해 안정적이고 책임 있는 프레임워크를 지속적으로 구축해야 함

3 미국의 기초 AI 역량 증진 및 확보

① (역량 강화 - 인재) AI 분야 경쟁력 강화를 위한 국내외 AI 인재 유치 및 육성(3.1.(b)~(c))

- (전문가 유치) 국무·국방·국토안보부는 AI·반도체 설계·생산 등 전문가 유치를 위한 모든 법적 권한 활용
- (인재 시장 분석) 180일 이내 경제자문위원회는 미국·해외 AI 인재시장 분석 준비
- (민간 경쟁력 평가 조정 및 정책 권고) 180일 이내 경제정책보좌관·국가경제위원회는 민간 AI 생태계 경쟁우위의 주요 원천 분석 및 위험 완화 정책 권고
- (절차 간소화) 90일 이내 국가안보보좌관은 해당 행정부처 및 기관을 소집하여 민감기술 분야 비자신청 행정 처리 간소화 조치 모색

② (역량 강화 - 컴퓨팅 인프라) 최첨단 AI 반도체 개발 및 AI 전용 컴퓨팅 인프라 구축(3.1.(d))

- (컴퓨팅 시설 구축) 국방부, 에너지부, 정보기관은 대규모 AI 적용을 고려한 첨단 컴퓨팅 시설 설계·구축
- (컴퓨팅 자원 제공) 국립과학재단(NSF)은 NAIRR(국가 인공지능 연구자원)을 통해 대학·비영리단체 등에 컴퓨팅 자원 및 데이터 등 제공
- (파일럿 프로젝트 착수) 180일 이내 에너지부는 AI 규모 훈련·미세조정·추론을 위한 연합 AI·데이터 소스 성능 및 효율성 평가 시범사업 착수
- (인프라 구축 지원) 백악관 비서실장실은 에너지부 및 기타 관련 기관과 협력하여 청정에너지 발전, 송전선, 고용량 광섬유 데이터 링크 등 인프라 구축 지원
- (투자 확대) 국무부, 국방부, 에너지부 등은 국내외 전략적 AI 기술 및 인접 분야에 대한 공공 투자 시행 및 민간 투자 장려

③ (해외 위협 대응) 미국 AI 생태계를 외국 정보 위협으로부터 보호(3.2)

- (위협 평가) 90일 이내 국가안보위원회(NSC)·국가정보국(ODNI)은 AI 생태계·반도체 분야 해외위협 식별·평가 개선하기 위한 권고안 수립
- (공급망 보호) 180일 이내 국가정보국은 협력을 통해 AI 공급망 주요 노드 식별 및 외국 행위자의 교란·손상 가능 경로를 파악하고 위험을 경감하기 위한 조치 이행
- (외국 행위자의 지적 재산 위협으로부터 보호) 기술이전·데이터 현지화 등 회색지대 방법을 통한 외국 행위자의 지적재산 획득 시도 경계
- (투자 심사) 외국인투자위원회는 해외 주체가 AI 교육 기술, 알고리즘 개선, 하드웨어 발전, 핵심 기술 산출물(CTAs*) 또는 강력한 AI 시스템을 만드는 방법에 대한 독점적 통찰력에 접근하는지 여부 고려

  * Critical Technical Artifacts

④ (AI 안전·보안·신뢰성 강화) AI의 긍정적 잠재력 실현을 위해 AI 위험을 평가하고 완화하기 위한 테스트 인프라를 선제적으로 구축(3.3)

- (안전성 테스트) 상무부는 NIST 내 AI 안전연구소를 통해 프론티어 AI 모델의 자발적인 사전 및 사후 공개 배포 안전 테스트 촉진

  ※ 사이버 보안, 생물안보, 화학무기, 시스템 자율성 등 분야별 위험 평가가 포함되며, 핵 위험은 제외(핵 위험 평가는 에너지부가 주도)
- (평가지원 조치 이행) 상무부는 AI 안전연구소를 통해 AI 시스템 평가를 지원하기 위해 아래 조치 이행
  · 180일 이내 최소 2개 이상 프론티어 AI 모델 예비 테스트 실시
  · 180일 이내 이중용도 기반모델 위험 평가 및 관리 지침 마련
  · 180일 이내 AI 시스템 과학·수학·코드생성·추론 분야 평가 벤치마크 개발
  · 270일 이내, 그리고 매년 AI 안전연구소는 국가안보보좌관을 통해 대통령에 AI 안전성 평가에 대한 보고서 제출
- (분야별 대응) AI 안전연구소를 통해 사이버·핵·방사능에 대한 부문별 평가 진행
  · 120일 이내 국가안보국은 AI 모델의 사이버 위협 평가 역량 개발
  · 120일 이내 에너지부는 핵·방사능 위험 평가 역량 개발
  · 210일 이내 에너지부, 국토안보부, AI 안전연구소는 생물·화학 위협 평가 로드맵 수립하여 국가안보보좌관에 공유

4 국가안보 목표 달성을 위한 AI 활용

① 효과적이고 책임있는 AI 활용 활성화(4.1)

- (인재확보) 국무부, 국방부, 법무부, 교육부, 국토안보부·정보기관은 책임있는 AI 가속화를 위한 인재 채용·유지 정책 및 전략 수정
- (교육강화) 120일 이내 국무·국방·법무·교육·국토안보부·정보기관은 각각 관리예산처(OMB)와 협의하여 AI 역량 제고를 위한 교육 및 훈련 마련
- (조달 시스템 개선) 국가안보 임무 수행에 AI를 사용하는 것을 가속화하기 위해 효과적인 조달 시스템 운영
  - · 30일 이내 국방부·국가정보국 주도 AI 조달·활용 실무그룹 구성
  - · 210일 이내 위 실무그룹은 연방 조달 규제 위원회(FARC)에 서면 권장사항 제공

    ※ 실무그룹은 AI 시스템의 안전, 보안 및 신뢰성을 측정하고 촉진하기 위한 객관적 지표 보장, 안전 위험을 완화하기 위한 적절한 점검 유지 등 수행
- (정책개선) 국방부·정보기관은 법무부와 협의하여 국제법과 국내법 등에 따라 AI 관련 법적·정책적 프레임워크 검토 및 개선
- (동맹국 협력) 150일 이내 국방부는 국무부·국가정보국과 협력하여 주요 동맹국과의 AI 및 AI 지원 자산의 공동 개발 및 공동 사용의 발전, 증대, 촉진의 타당성 평가

② AI 거버넌스 및 위험 관리 강화(4.2)

- (기본 방향) AI가 국가안보 임무를 지원할 때 인권, 시민권, 시민적 자유, 프라이버시 및 안전을 보호하고 군사작전에서 책임있는 인간 지휘체계 유지 필요
- (위험 관리) 관련 기관 책임자는 AI 개발 및 사용으로 인한 위험을 모니터링, 평가, 완화해야 함

  ※ 신체적 안전에 대한 위험, 개인정보 침해, 차별 및 편견, 부적절한 사용, 책임감 부족, 데이터 유출, 성능 저하, 고의적 조작 및 오용 등 9개 분야 위험
- (프레임워크 구축) AI 프레임워크는 국가안전보장회의 부의장 위원회에서 승인을 받아야하며 프로세스를 통해 위 명시된 위험과 AI 프레임워크에서 다루는 기타 주제를 해결하기 위해 조정이 필요한지 여부 결정

  ※ AI 프레임워크는 OMB의 '기관의 인공지능 사용에 대한 거버넌스, 혁신 및 위험관리 발전에 관한 각서(M-24-10)' 또는 후속 OMB 정책에 대한 국가안보 중심의 대응책 역할
  - · (조직 구성) 각 기관 AI 최고 책임자 지정 및 AI 거버넌스 위원회 구성
  - · (위험 지침) 고위험 AI 활동 식별 및 최소 위험관리 지침 마련
  - · (이행 관리) 교육·훈련 프로그램 및 책임성 확보 체계 구축, 고위험 AI 활용·시스템 연례 조사·보고·내부고발 보호제도 운영, 국가안보전략 관리자는 150일 이내 최소한의 AI 사이버보안 지침 발표

- (국가안보시스템 지침) 국가안보시스템(NSS)에서 AI 사용에 관한 구체적 지침 제공
  · 180일 이내 국가안보시스템의 일부로 AI를 사용하는 모든 부처·기관장은 AI 거버넌스·위험관리 지침을 발행하고 연례 조정회의 개최

5 국제 AI 거버넌스

○ (국제협력 강화) 동맹국 및 파트너와 AI 거버넌스 관련 협력 확대

- (전략 수립) 국무부는 120일 이내 국방부, 상무부, 국토안보부, UN 대표부, 국제개발처와 협력하여 국제 AI 거버넌스 규범 발전 전략 수립
- (규범 개발) 미국 국내 정책과 일관된 국제 규범·표준 개발 추진
  ※ G7 AI 시스템 개발자 행동강령, OECD AI 원칙, UN 총회 결의안 등과 정합성 확보

6 AI 정책의 효과적인 조정·실행·보고체계 구축

○ (조정체계) 효과적이고 책임있는 AI 도입을 위해 긴밀하고 조율된 방식으로 협력

- 45일 이내 국무부, 국방부, 법무부, 교육부 등의 최고 AI 책임자와 적절한 기술 담당자가 AI 국가안보 조정 그룹 구성
- 90일 이내 조정그룹 내 모든 기관의 고위 AI 관계자로 구성된 국가 안보 AI 인재위원회 설립
  ※ AI 인재 요구 사항 표준화, 우선순위 지정 및 해결하고 국가안보 목적의 AI 및 AI 활용 인재 유치, 채용, 개발 및 유지를 위한 정부 차원 대응
- 270일 이내 각 부처는 본 각서의 조항별 임무에 대한 설명과 추가 조치 계획에 대한 최초 보고서를 국가안보보좌관을 통해 대통령에 제출
- 365일 이내, 그리고 향후 5년간 조정그룹은 국가 안보를 목적으로 하는 AI 노력과 시스템의 통합 및 상호운용성 관련 공동 보고서를 국가안보보좌관에 작성·제출

7 정의

○ (AI 안전성) AI 모델의 악의적 사용, 오용, 실패, 사고 및 의도치 않은 행동으로 인해 개인과 사회에 발생할 수 있는 피해 가능성을 최소화하고 완화하기 위한 메커니즘을 의미

○ (AI 보안) AI 시스템(훈련 데이터, 모델, 기능 및 수명주기 포함)을 사이버 및 물리적 공격, 도난, 손상으로부터 보호하기 위한 일련의 관행을 의미

○ (대상 기관) 정보기관과 국가안보시스템의 구성요소로 AI를 사용하는 모든 기관(대통령 집행실 제외)을 의미

○ (핵심 기술 산출물(CTAs)) 모델 개발자가 아닌 다른 사람이 보유할 경우 해당 모델의 기능을 재현, 획득 또는 사용하는 비용을 실질적으로 낮출 수 있는 정보를 의미

- 현재 AI 산업의 기술 패러다임에서는 훈련된 AI 시스템의 모델 가중치가 CTAs를 구성하며, 경우에 따라 관련 훈련 데이터와 코드도 포함

○ (프론티어 AI 모델) 널리 인정되는 공개 벤치마크 또는 추론, 과학 및 전반적 능력에 대한 유사한 평가로 측정된 최첨단 성능에 근접한 범용 AI 시스템을 의미

○ (정보기관(IC)) 美 연방법 제50장(50 U.S.C.) 3003*에 규정된 의미

* 중앙정보국(CIA), 국방정보국(DIA), 국가안보국(NSA), 국가정찰국(NRO), 국가지리정보국(NGA), 연방수사국(FBI)의 정보 및 국가안보 부서, 국무부 정보조사국(INR), 재무부, 에너지부, 국토안보부 등 각 부처의 정보 부서 등

○ (오픈웨이트 모델) 일반적으로 공개 릴리스를 통해 널리 사용 가능한 가중치를 가진 모델을 의미

○ (미국 정부) 美 연방법 제44장(44 U.S.C.) 3502(1)에 정의된 모든 기관을 의미

* 국무부·국방부·재무부 등 행정부 부처, 환경보호청·연방통신위원회 등 독립행정기관·군사 관련 부처 등

8 일반 조항

○ 본 각서의 어떠한 내용도 행정부서나 기관 또는 그 장에게 법률로 부여된 권한, 예산/행정/입법 제안과 관련된 관리예산처장의 기능을 손상시키거나 영향을 미치는 것으로 해석되지 않음

○ 본 각서는 적용 가능한 법률 및 예산의 가용성에 따라 시행되어야 함

○ 본 각서는 미국, 그 부서, 기관 또는 단체, 그 임원, 직원 또는 대리인 또는 기타 어떤 사람에 대해서도 법률이나 형평법상 집행 가능한 실체적 또는 절차적 권리나 혜택을 창출하지 않으며, 그러한 의도도 없음

| 참고 | 국가 안보의 AI 거버넌스 및 위험 관리 발전을 위한 프레임워크 주요내용

■ 바이든 정부는 각서와 '국가 안보의 AI 거버넌스 및 위험 관리 발전을 위한 프레임워크'를 동시 발표

■ (목적) 미국 정부가 인권, 민권, 시민의 자유, 프라이버시 및 안전 보장을 위한 조치를 지속적으로 취하며, 군사작전 중 책임 있는 AI 사용 등을 보장하기 위한 목적으로 AI 프레임워크* 수립

* 국가 안보에서 AI 거버넌스 및 위험 관리를 발전시키기 위한 프레임워크(이하 'AI 프레임워크')

※ 군사적 맥락의 AI 사용은 '인공지능과 자율성의 책임있는 군사적 사용에 관한 정치 선언('23.11.9)'에 명시된 원칙과 조치 준수

■ (범위) AI 프레임워크는 국가 AI 안보 각서 4.2항(AI 거버넌스 및 위험관리 강화)에 명시된 요구사항 기반으로 수립

- (OMB 각서) 관리예산처(OMB)의 각서(M-24-10)*를 보완하며, 연방기관의 모든 AI 사용은 OMB 각서와 후속 정책 및 본 프레임워크 적용

  * 기관의 인공지능 사용에 대한 거버넌스, 혁신 및 위험관리 발전에 관한 각서

- (국가안보시스템) AI가 국가 안보 시스템 구성 요소로 사용될 때 적용
- (정부 AI) 미국 정부에 의해, 또는 정부를 대신하여 개발·사용·조달되는 신규 및 기존 AI에 모두 적용

■ (업데이트) 국가안보위원회(NSC) 부위원회 회의에 따라 업데이트 진행

※ 국가안보각서-2('21.2.4)(국가안보위원회 시스템 갱신)에 설명된 절차 준수

■ 주요내용

- 금지되고 영향력이 큰 AI 사용 사례와 연방 직원에게 영향을 미치는 AI 사용 사례 식별 및 규제
- 영향력이 큰 것으로 식별된 AI 범주에 대해 충분히 강력한 최소 위험관리 방법 수립
- 영향력이 큰 AI 사용을 분류하고 모니터링
- 효과적인 교육 및 책임 매커니즘 보장에 관한 내용 포함

| 금지되고 영향력이 큰 AI 사용 사례 |

금지된 AI 사용	■ 헌법상 권리 행사 기반 개인 프로파일링/추적 ■ 표현/법적 조언의 자유 억압 ■ 차별(인종,성별,종교 등) 목적 사용 ■ 개인 감정상태 추론(허가된 경우 제외) ■ 생체인식 데이터만으로 개인 정체성 판단	■ 적절한 감독 없는 전투 피해 추정 ■ 이민/망명 최종 결정 ■ AI 단독 기반 정보분석 보고 ■ 핵무기 관련 결정에서 인간 배제
고영향 AI 사용 *특별 안전장치 필요	■ 생체인식 기반 실시간 개인 추적 ■ 테러리스트/안보위협 인물 분류 ■ 이민/망명 관련 결정	■ 위험물질/시스템 관리 ■ 자동 코드 작성/수정 ■ 단독 정보분석 생산연방 인사 관련 AI
연방 인사에 영향을 미치는 AI 사용	■ 채용 결정/임금 책정 ■ 승진/강등/해고 결정	■ 직무성과/건강상태 평가추가 규제

분석 및 시사점

○ 미국은 AI NSM 발표를 계기로 AI 기술을 핵무기, 우주기술과 같은 **국가 전략 자산**으로 간주하고 **정부의 핵심 우선순위**로 AI를 설정하여 통제하고 관리하기 시작

- 그간 민간 기업에 대한 단순 감독(자발적 규제) 수준을 넘어 국가 차원에서 AI 인재 유치 등 AI 개발 지원과 잠재적 위협을 통제*하기 위해 정부 역량을 총동원

* AI 시스템의 군사적, 비인도적 오용 방지 및 AI 공급망 파괴 대응 등

○ 동시에 'AI 굴기'를 추진 중인 중국에 대응하기 위한 조치로서, 경쟁국을 견제하는 동시에 동맹국과의 협력 강화를 통해 **AI 분야에서 미국이 독보적 우위를 점하겠다는 의지**로 해석될 수 있음

※ 중국은 '30년까지 AI 핵심산업 규모 1조 위안(약 195조원), 관련산업 규모 10조 위안(1950조원) 달성 목표 정책 추진(SCMP, '24.9.9)

※ 중국과학기술정보연구소(ISTIC)가 베이징대와 함께 발표한 '2023 AI 글로벌 혁신 지수 보고'에 따르면 상급 학술지에 실린 AI 논문 점유율(중국 36.7%, 미국 22.6%)에서 중국이 미국을 앞선 것으로 분석(한국경제, '24.7.7)

- 미국이 동맹국, 파트너와의 AI 협력을 강조하는 가운데 한국도 미국과 협력을 통해 첨단 AI 기술과 인프라에 대한 접근성을 제고할 기회 마련 가능
- 반면, 미국의 AI 우위 강화는 동맹국들의 상대적 자율성을 약화시킬 수 있는 가능성도 존재하므로 미국과 협력을 통해 이익을 최대화하되 기술 자주권을 확보하기 위한 전략 마련 필요(Just Security, '24.10.25)

○ **(인재)** AI 인재 확보를 단순히 산업 경쟁력의 문제가 아닌 **국가안보**의 **우선순위**로 **격상***시켜 **국내 인재 육성, 해외 인재**의 신속한 유치, **정부 AI 인재 관리 혁신**을 모색

* 주요 이행 내용의 첫 번째로 AI 인재와 이민 정책을 포함

- AI 및 반도체 설계·생산 등 연관 분야의 해외 인재를 신속하게 유치를 위한 가능한 모든 법적 권한을 활용하는 등 정부의 적극적 개입을 강조
- AI 전문가 채용을 위한 인사관리 제도 개선, 내부 인력에 대한 교육 훈련 강화, 산업-정부 간 인재 교류 활성화하고 글로벌 AI 인재 동향을 지속적으로 모니터링하는 체계를 구축

○ **(인프라)** 美 행정부는 전력, 컴퓨팅 자원, 데이터센터 등 **물리적 인프라**를 **확대**할 수 있도록 **조정 예정**으로 예산 권한의 한계로 역할은 제한적*이나 백악관 비서실장에 임무를 부여하며 우선 과제로 설정

* 직접 예산 확대는 의회, 주지방 정부에 권한이 있으므로 행정부는 인프라 건설, 청정에너지 생산 등 주변 자산에 대한 허가, 승인 등 절차 간소화 및 인센티브 부여 등을 위한 조정 역할 예정

○ **(공급망)** 미국의 대중 견제 심화로 기술 공급망도 재편되고 있는 가운데 한국 역시 지속 가능한 AI 공급망을 확보하기 위해 전략적 파트너 국가들과 협력을 유지하는 노력이 필요할 것이라 사료됨

- 미국은 AI 공급망의 주요 취약점 식별, 잠재적 파괴 경로 파악, 위험 경감 조치 등 AI 생태계 전반의 공급망 안정성 확보를 위한 세부 전략을 세울 예정으로 선제적 대응을 위한 대비책 마련

○ **(안전성)** 미국은 **AI 안전성**을 AI 혁신을 저해하거나 늦추는 규제가 아닌 오히려 **더 빠른 혁신과 도입을 가능하게 하는 촉진제**로써 명확한 안전 기준과 거버넌스 체계가 오히려 정부 기관들의 AI도입을 가속화할 수 있다고 인식

※ 제이크 설리번 국가안보보좌관은 "*보안과 신뢰성을 확보하는 것은 우리의 발걸음을 늦추는 것이 아니라 오히려 더 빠른 진전을 가능하게 할 것…안전성과 신뢰성에 대한 확신이 없으면 새로운 기술을 실험하고, 도입하고, 활용하는 데 더욱 조심스러워질 수밖에 없는데, 오늘날의 전략적 환경에서 우리는 그럴 만한 여유가 없다*"고 언급

- 범정부 차원의 AI 시스템 사이버보안 최소 기준, 보안 취약점 평가·대응, 위험정보 공유 등 종합적 보안 체계를 구축하며 지속적으로 관리하고 업데이트 중

○ **(AI 활용 및 AISI) 프론티어 AI**를 **국가 경쟁력의 결정적 요소**이자 **국가안보를 위한 주요한 도구**로 **인식**하며 **정부 기관의 적극적 활용을 장려**하며 이에 대한 **안전성, 보안성, 신뢰성 보장을 AI 안전연구소(AISI)***의 **주요 임무**로 강조

* 미국은 상무성 산하 국가표준기술연구소(NIST) 내 AI 안전연구소를 설립('23.12)하고 200여개 기관들이 참여하는 'AI 안전 컨소시움'을 발족('24.2월), AI의 안전한 개발과 배포를 위한 표준 연구 추진 중

- AI NSM의 안전 관련 섹션은 대부분이 AI 안전연구소(AISI)의 업무에 초점이 맞춰져 있으며 본 내용이 AI 안전연구소의 구체적인 업무가 될 것 예정

〈 AI 국가보안각서 내 AI 안전연구소(AISI)의 주요 역할 〉

- AI 테스트 및 평가 활동과 관련하여 AISI를 민간 부문 AI 기업들의 주요 연락 담당 기관으로 지정
- AISI에 180일 이내에 적어도 2개의 프론티어 AI 모델에 대해 공개 배포나 출시 이전에 국가안보에 위협이 될 수 있는 능력을 평가하기 위한 자발적 예비 테스트를 추진하도록 지시
- AISI에 180일 이내에 이중용도 기반 모델에서 발생하는 안전성, 보안성, 신뢰성에 대한 위험을 테스트, 평가, 관리하는 방법에 대한 AI 개발자 지침을 발표하도록 지시
- AISI와 국가안보국의 AI보안센터가 AI 모델이 공격적 사이버 위협을 탐지, 생성, 악화시키는 능력에 대한 신속한 체계적 기밀 테스트를 수행할 수 있는 능력을 개발하기 위해 긴밀히 협력하도록 지시

○ **(정부-민간 협력)** 정부의 직접 개발보다 **민간 주도의 AI 혁신**을 **인정**하고 **민간 AI 기업 보호·육성 등 AI 생태계**를 **지원**하는 정부의 역할을 강조하며 국가안보와 기술 혁신의 상생 모델을 제시

※ 특히 상업용 AI 기업들을 보호하고 그들의 민감한 기술을 보호하는데 관여할 예정

○ **(동맹국 및 파트너와의 협력)** AI NSM에 따라 '국제 AI 거버넌스 규범 발전 전략 수립'(5.(c).(i)) 후 **자국의 정책**에 **맞춰 글로벌 규범**을 **적극적**으로 **주도**하고 동맹국 및 파트너와의 협력을 강화해 나갈 것으로 예상

- 미국은 국제 파트너들과의 협력을 통해 AI 거버넌스의 국제적 발전을 추구하며, 이를 통해 기술 혁신과 함께 글로벌 규범 형성에서도 리더십을 확보하고자 함

- 목적 조항(Sec.2), 글로벌 거버넌스 조항(Sec.5)에서는 **동맹국 및 파트너와의 협력을 강조하는 한편, AI 활용 조항(Sec.4)에서**는 AI 역량의 공동 개발에 대한 투자, AI 공동 개발·사용의 타당성 평가에 대한 협력 주체로 **'선별적' 동맹국 및 파트너**(select allies and partners)**로 그 범위를 제한**

※ 이는 미국이 동맹국 및 파트너와의 협력에 대해 거버넌스 차원보다 AI 활용 측면에서 전략적 접근을 취하고 있다는 것을 시사

[표 2] 동맹국 및 파트너 명시 조항 예시

동맹국 및 파트너 명시 조항	'선별적(select)' 동맹국 및 파트너 명시 조항
■ **2(c)** : 안전하고 신뢰할 수 있는 AI 개발 및 사용을 촉진하는 국제 AI 거버넌스를 발전시키기 위한 프레임워크 구축을 위해 다양한 동맹국 및 파트너와 협력해야 한다.	■ **4.1(h)** : 2022년 국가 안보 전략 및 후속 전략에 따라, 미국 정부는 선별된 동맹국 및 파트너와 AI 성능의 공동 개발·배포에 투자하고 이를 적극적으로 지원해야 한다.
■ **5(c)(i)** : 국무부는 (중략) 국제 AI 거버넌스 규범의 발전을 위한 전략을 수립한다. 이 전략은 양자 및 다자간 참여와 동맹국 및 파트너와의 관계를 다루어야 한다. 또한, 경쟁국과의 관계에 대한 지침을 포함해야 하며, 유엔 및 G7과 같은 국제기구 및 기술 기구와의 협력에 대한 접근 방식을 설명한다.	■ **4.1(h)(i)** : 국방부는 선별된 동맹국 및 파트너와 함께 AI 및 AI 지원 자산의 공동 개발 및 공동 사용의 발전, 증대, 촉진의 타당성을 평가해야 한다.

○ **(거버넌스)** 각 기관의 AI최고책임자 지정(4.2,(ii), (A)), AI최고책임자로 구성되는 AI 국가안보 조정 그룹(6.(b),(ⅱ)), AI거버넌스위원회 설치(4.2,(ii),(B))까지 **체계적이고 다층적인 거버넌스**를 통해 **정부의 AI 역량**을 **결집**

[표 3] 전문가들의 AI NSM에 대한 평가(Just Security, '24.10.25)

구분	전문가	평가 내용
1	Suresh Venkatasubramanian (브라운대학 교수, 전 백악관 과학기술정책실 차관보)	■ NSM이 AI의 책임있는 사용, 시민의 자유와 권리 보호에 대해 강력한 언어를 사용한 것은 고무적 ■ 하지만 예외조항들이 많아 실제로 얼마나 효과적일지는 시간이 지나봐야 알 수 있음 ■ Framework 문서가 더 희망적이라고 평가 (쉽게 업데이트 가능한 구조, 금지된 사용 사례 명시 등) ■ 전반적으로 예상보다 강력한 보호조치를 포함했다고 긍정적 평가
2	Ashley Deeks (버지니아 로스쿨 교수, 전 백악관 법률자문)	■ 민주적 가치와 국가안보 AI의 도전과제를 진지하게 다룬 점을 높이 평가 ■ 하지만 외부 감독 메커니즘이 부재한 점을 지적 ■ 의회에 대한 보고 의무가 없는 것이 문제 ■ "high impact" AI 활동에 대한 기관간 검토 프로세스 필요성 강조
3	Brianna Rosen (저스트 시큐리티 선임연구원, 옥스퍼드 대학교 전략 및 정책 연구원)	■ AI 위험관리를 위한 최초의 정부 차원 프레임워크라는 점에서 의미있는 진전 ■ 하지만 정보기관의 면제 프로세스에 대한 우려 제기 ■ 외부 감독 기구 부재 문제 지적 ■ 투명성과 책임성 메커니즘이 부족하다고 평가
4	Thompson Chengeta (리버풀 존 무어스 대학교 국제법 및 인공지능 교수)	■ 국제법적 관점에서 평가 ■ 국제법의 모든 관련 분야를 균형있게 고려해야 할 필요성 강조 ■ 차별 '최소화'가 아닌 '제거'를 목표로 해야 한다고 지적 ■ 법적 구속력 있는 국제조약 필요성 강조
5	Keith Dear (후지쯔 인지·첨단기술센터 이사, 전 영국 총리 국방 현대화 및 통합 검토 전문가 고문)	■ 미국의 AI 리더십 확보 의지가 명확히 드러난 점을 강조 ■ 동맹국들에게는 안심과 우려를 동시에 주는 문서라고 평가 ■ 중국보다 미국이 AGI/ASI를 선점하는 것이 민주주의 국가들에게 더 나은 시나리오 ■ 하지만 미국의 우위가 동맹국들의 상대적 힘을 약화시킬 수 있다고 우려
6	Faiza Patel (브레넌 센터 자유와 국가안보 프로그램 선임 디렉터)	■ 국가안보 AI 문제를 다룬 것 자체를 긍정적으로 평가 ■ 하지만 대부분의 작업이 비공개로 이뤄질 것을 우려 ■ 내부 감독 메커니즘에 과도하게 의존한다고 지적 ■ 새 행정부 출범 시 지속될지 불확실하다고 우려
7	Patrick Toomey (미국 시민 자유연합(ACLU) 국가 안보 프로젝트 부국장)	■ 일부 긍정적 조치에도 불구하고 충분치 않다고 평가 ■ 독립적 감독, 투명성, 통지, 시정 조치 부분이 크게 미흡 ■ 국가안보기관들의 자체 감독은 위험할 수 있다고 지적 ■ 피해자들을 위한 구제 메커니즘 부재 비판
8	Brandon Pugh (R Street Institute 사이버보안 및 신종 위협 정책 책임자)	■ AI를 국가안보에 활용하는 것을 긍정적으로 평가 ■ 규제와 혁신 사이의 균형 필요성 강조 ■ 중국 등 경쟁국들이 미국의 제한을 따르지 않을 것이라는 우려 ■ 정권 교체 시 정책 지속성에 대한 우려 제기
9	Bill Drexel (신미국안보센터 기술 및 국가 안보 프로그램 펠로우)	■ 국내 컴퓨팅 능력 강화, AI 도입, 사이버보안 강화 등을 긍정적으로 평가 ■ 하지만 개발도상국의 AI 생태계 구축 지원이 부족하다고 지적 ■ 프론티어 모델 개발에 과도하게 초점을 맞추고 있다고 비판 ■ 국가안보를 위한 다양한 AI 도구 개발의 균형이 필요하다고 제안

[참고1] 주요 내용의 이행주체 및 기한

구분	주요 내용	이행주체 및 기한
[섹션 3] 3.1 미국 AI 개발의 진보, 혁신 및 경쟁 촉진	■ 신뢰할 수 있는 데이터가 있는 범위 내에서 미국 및 해외의 AI 인재 시장에 대한 분석 준비	경제자문위원회 의장 /180일 이내
	■ 미국 민간 부문 AI 생태계의 상대적 경쟁 우위, 미국 민간 부문 경쟁 우위의 주요 원천 및 그러한 위치에 대한 가능한 위험에 대한 경제적 평가를 조정하고 이를 완화하기 위한 정책 권고 〈포함내용〉 (1) AI 관련 활동에 중요한 칩의 설계, 제조 및 패키징 (2) 자본의 가용성 (3) AI 관련 분야에 고도로 숙련된 인력의 가용성 (4) 컴퓨팅 자원 및 관련 전기 요구 사항 (5) 프론티어 AI 모델 개발에 필요한 규모의 자본 및 데이터 자원을 보유한 기술 플랫폼 또는 기관 및 기타 가능한 요인을 포함한 영역이 포함	대통령 경제정책 보좌관 및 국가경제위원회 국장 /180일 이내
	■ 해당 행정부처 및 기관(단체)을 소집하여 민감한 기술을 다루는 모든 비자 신청자에 대한 행정 처리 작업의 우선순위를 정하고 간소화하는 조치 모색	대통령 국가안보보좌관 /90일 이내
	■ 프론티어 AI 규모의 훈련, 미세 조정 및 추론을 위한 연합 AI 및 데이터 소스의 성능과 효율성을 평가하기 위한 파일럿 프로젝트 시작	에너지부(DOE) /180일 이내
[섹션 3] 3.2 외국의 정보 위협으로부터 미국 AI 보호	■ 대통령의 정보 우선순위와 국가정보 우선순위 프레임워크를 검토 – 우선순위가 미국 AI 생태계와 반도체 설계 및 생산과 같은 밀접하게 관련된 지원 부문에 대한 외국 정보 위협의 식별 및 평가를 개선하도록 권고	국가안보회의(NSC) 직원 및 국가정보국장실(ODNI) /90일 이내
	■ AI 공급망에서 중요한 노드를 식별하고 이러한 노드가 외국 행위자에 의해 중단되거나 손상될 수 있는 경로 목록을 개발 ※ 국방부, 법무부(DOJ), 상무부, DOE, DHS 및 기타 적절한 IC와 협력	국가정보국장실(ODNI) /180일 이내
[섹션 3] 3.3 AI 안전, 보안 및 신뢰성에 대한 위험 관리	■ 국가 안보에 위협이 될 수 있는 기능을 평가하기 위해 공개 배포 또는 출시 전에 최소 2 개의 프론티어 AI 모델에 대한 자발적인 예비 테스트를 추진	상무부 산하 국립표준기술연구원(NIST)의 A안전연구소(AISI) /180일 이내
	■ 이중 사용 기반 모델에서 발생하는 안전, 보안 및 신뢰성에 대한 위험을 테스트, 평가 및 관리법에 대해 AI 개발자를 위한 지침 발행 〈포함내용〉 (1) AI 모델이 생화학 무기 개발 또는 공격적인 사이버 작전의 자동화를 가능하게 할 수 있는 위험과 관련된 역량을 측정하는 방법 (2) 개인을 괴롭히거나 사칭하기 위한 모델 오용과 같은 사회적 위험에 대처하는 방법 (3) 모델의 악의적이거나 부적절한 사용을 방지하기 위한 완화 조치를 개발하는 방법 (4) 안전 및 보안 완화 조치의 효과 테스트 방법	상무부 산하 NIST의 AISI /180일 이내

구분	주요 내용	이행주체 및 기한
	(5) 개발 및 배포 수명 주기(개발 전, 개발, 배포/출시) 전반에 걸쳐 위험 관리 관행을 적용하는 방법	
	■ 과학, 수학, 코드 생성 및 일반 추론에서 AI 시스템의 능력과 한계를 평가하는 벤치마크 또는 기타 방법과 국가 안보 및 공공 안전에 영향을 미칠 수 있는 범용 능력을 평가하는 데 관련이 있다고 판단하는 기타 범주의 활동을 개발하거나 권장	상무부 산하 NIST의 AISI /180일 이내
	■ 이하 내용에 대한 최소 1건의 보고서 제공 〈보고서 내용〉 (1) AISI가 수행했거나 공유한 프론티어 AI 모델에 대한 AI 안전성 평가 결과 요약 (2) 평가에서 확인된 문제를 해결하기 위해 위험 완화가 필요하다고 판단했는지 여부에 대한 요약과 완화 효과에 대한 결론 (3) 평가를 알리는 데 사용된 과학 기반 도구 및 방법의 적절성에 대한 요약	상무부 산하 NIST의 AISI/270일 이내 (최소 매년 주기)
	■ 기밀 및 통제 정보에 대한 해당 보호 조치에 따라 완료 후 NIST AISI에 보고	AI 시스템의 안전성 테스트 및 평가를 수행하거나 자금을 지원하는 모든 기관 /조치 완료 후 30일 내
	■ AISI와 협력하여 공격적인 사이버 위협을 탐지, 생성 및/또는 악화시키는 AI 모델의 능력에 대한 체계적인 기밀 테스트를 신속하게 수행할 수 있는 역량을 개발	국가안보국(NSA) (AI보안센터를 통해 보고) 120일 이내
	■ 국가핵안보국(NNSA)을 통해 그리고 AISI 및 NSA와 긴밀히 협력하여 핵 및 방사능 위험을 생성하거나 악화시키는 AI 모델의 능력에 대한 신속한 체계적 테스트를 수행할 수 있는 역량을 개발하기 위해 노력	에너지부 /120일 이내
	■ DOE는 NNSA(국가핵안보국)를 통해 AI 모델의 핵 및 방사능 위험을 평가할 수 있는 신속한 체계적 테스트 기능을 개발 - 테스트 기능을 활용해 프론티어 AI 모델이 제공된 후 30일 내 해당 모델의 방사능 및 핵관련 지식, 관련 능력, 잠재정 영향에 대한 초기 평가 완료 〈테스트 기능〉 (1) AI모델의 핵 및 방사능 위험을 평가할 수 있는 신속한 체계적 테스트 기능 개발 (2) 기밀/비기밀 테스트를 실행할 수 있는 인프라 포함 (3) 자동화된 평가, 인간 주도 레드팀 테스트 인터페이스, 정부/오픈웨이트/독점 모델의 신속하고 안전한 전송을 위한 기술적/법적 도구 포함	에너지부 /180일 이내
	■ 이하를 포함하는 최소 한 차례의 평가를 APNSA를 통해 대통령에게 제출 〈포함내용〉	에너지부 /270일 이내 (최소 매년 주기)

구분	주요 내용	이행주체 및 기한
	(1) 방사능 및 핵 위험에 대한 각 AI 모델 평가 결과의 간결한 요약 (2) 원자력법상 기밀정보 보호를 위한 시정조치 권고사항 검토 (3) 평가를 알리는 데 사용된 과학 기반 도구 및 방법의 적절성에 관한 서술	
	■ 국방부 및 기타 관련 기관과 협의하여 고의적인 화학 및 생물학적 위협을 생성하거나 악화시키는 첨단 AI 모델의 능력에 대한 향후 기밀 평가를 위한 로드맵을 개발하여 APNSA와 공유	에너지부, 국토안보부, AISI /210일 이내
	■ 고의적인 화학 및 생물학적 위협을 발생시키거나 악화시키는 첨단 AI 모델의 능력에 대한 기밀 테스트를 수행할 수 있는 전문 지식, 인프라 및 시설을 제공하기 위한 시범 프로젝트 수립	에너지부 /270일 이내
	■ 생물안전 및 생물보안 강화를 위해 고성능 컴퓨팅 자원과 AI 시스템을 활용하려는 노력을 적절히 지원 〈포함내용〉 (1) 인실리코 화학 및 생물학적 연구 및 기술 스크리닝을 위한 도구 개발 (2) 핵산 합성 스크리닝을 위한 알고리즘 생성 (3) 새로운 생명공학을 위한 고수준 보안 보장 소프트웨어 기반 구축 (4) 클라우드 랩 및 바이오 파운드리의 전체 주문 또는 데이터 스트림의 선별 (5) 의료 대책과 같은 위험 완화 전략 개발	국방부, 보건복지부, 에너지부(국립 연구소 포함), 국토안보부, 국립과학재단(NSF) 및 생물학적 및 화학적 데이터에 대해 실질적으로 훈련된 AI 시스템 개발을 추진하는 기타 기관 /240일 이내
	■ 학술 연구 기관 및 과학 출판사와 함께 AI를 사용하고 오용되어 해를 끼칠 수 있는 지식, 정보, 기술, 제품의 생산에 기여할 수 있는 것을 포함하여 전산 생물학적 및 화학 모델, 데이터 세트 및 접근법을 출판하기 위한 자발적인 모범 사례 및 표준을 개발하기 위해 노력 ※ 국방부, 상무부(NIST 내 AISI를 통해 대행), 보건복지부, DOE, 과학기술정책실(OSTP) 및 기타 관련 기관과 협력	NSF /180일 이내
	■ 인실리코 생물학 및 화학 연구의 혜택을 증진하고 관련 위험을 완화하는 지침 개발	과학기술정책실(OSTP), 국가안전보장회의(NSC) 직원, 팬데믹 대비 및 대응 정책실 /540일 이내
[섹션 4] 4.1 효과적이고 책임감있는 AI 사용 활성화	■ 각각 관리예산처(OMB)와 협의하여 교육 및 기술 기반 채용을 포함할 수 있는 이니셔티브를 통해 해당 인력의 AI 역량을 높이기 위한 교육 및 훈련 기회를 파악	국무부, 국방부, 법무부, 교육부, 국토안보부 및 IC /120일 이내
	■ AI 조달과 관련된 문제를 해결하기 위한 실무 그룹 설립	국방부, 국가정보국장실(ODNI) /30일 이내
	■ AI 조달에 대한 기존 규정 및 지침 변경에 관한 서면 권고안을 연방획득규제위원회(FARC)에 제공 〈포함내용〉 (1) AI 시스템의 안전성, 보안성, 신뢰성을 측정하고 홍보하기 위한 객관적인 지표 확보 (2) 안전 위험을 완화하기 위해 적절한 점검을 유지하면서 연방조달규정에 따라 AI의 획득 및 조달 프로세스를 가속화	AI 조달 실무그룹 /210일 이내

구분	주요 내용	이행주체 및 기한
	(3) 계약 경험이 없는 기업도 관련 계약을 위해 의미 있는 경쟁을 할 수 있도록 절차를 간소화하여 미국 정부가 다양한 AI 시스템에 접근하고 AI 시장이 경쟁력을 갖출 수 있도록 함 (4) 상호운용성을 촉진하는 요구사항을 포함하고 제안을 평가할 때 공급업체의 기술 역량을 우선시하는 등, 강력한 참여를 장려하고 정부에 최고의 가치를 달성할 수 있도록 경쟁을 구조화 (5) 관련 기관에서 가능한 한 최대한, 그리고 적절하게 AI의 공동 사용을 수용 (6) 특정 권한과 임무를 가진 기관이 적절하고 필요한 경우 다른 정책을 시행할 수 있도록 보장	
	■ 국무부 및 ODNI와 협력하여 일부 동맹국 및 파트너와 함께 AI 및 AI 지원 자산의 공동 개발 및 공동 사용의 발전, 증대, 촉진의 타당성을 평가 〈포함내용〉 (1) 공동 개발 또는 공동 배치가 가능할 수 있는 외국의 잠재적 목록 (2) 잠재적 아웃리치를 위한 양자 및 다자 포럼 목록 (3) 잠재적인 공동 개발 및 공동 배포 개념 (4) 공동 개발한 AI 기능에 대한 분류에 적합한 테스트 수단 제안 (5) 향후 AI 기능의 공동 개발 및 공동 배포를 위한 기초로 사용하기 위한 기존 프로그램, 계약 또는 약정에 대한 고려 사항	국방부 /150일 이내
[섹션 4] 4.2 AI 거버넌스 및 위험관리 강화	■ NSS의 구성 요소로 사용되는 AI에 대한 최소한의 사이버 보안 지침 및/또는 방향을 발표	국가안전국(NSS) 국가 관리자 /150일 이내
	■ NSS의 AI 거버넌스와 위험 관리에 관한 지침을 해당 구성요소/하부기관에 발행하거나 업데이트	국무부, 재무부, 국방부, 법무부, 상무부, 교육부, 국방부, ODNI(18개 IC 요소를 대행), 기타 NSS의 일부로 AI를 사용하는 모든 해당 기관의 장(부서장) /180일 이내
[섹션 5] 안정적이고 책임감 있으며 전 세계적으로 유익한 국제 AI 거버넌스 환경 조성	■ 국방부, 상무부, 국토안보부, 유엔주재 미국대표부(USUN), 미국국제개발처(USAID)와 협력하여 안전하고 안전하며 신뢰할 수 있는 AI 및 인권, 민권, 시민의 자유, 프라이버시를 포함한 민주적 가치에 부합하는 국제 AI 거버넌스 규범의 발전을 위한 전략을 수립 〈전략 방향〉 (1) 미국의 정책 및 기존 노력과 일치하는 국제적으로 공유되는 정의, 규범, 기대치 및 표준을 개발하고 홍보하여 전 세계에서 안전하고 안전하며 신뢰할 수 있는 AI 개발 및 사용을 촉진 (2) 민주주의 가치에 따라 관련 국제법을 준수하여 국가 안보 상황에서 책임감 있고 윤리적인 AI 사용을 촉진	국무부 /120일 이내

[참고2] 국가 안보의 AI 거버넌스 및 위험 관리 발전을 위한 프레임워크 세부내용

Ⅰ. AI 사용 규제	
금지된 AI 사용	■ 헌법상 권리 행사 기반 개인 프로파일링/추적 ■ 표현/법적 조언의 자유 억압 ■ 차별(인종,성별,종교 등) 목적 사용 ■ 개인 감정상태 추론(허가된 경우 제외) ■ 생체인식 데이터만으로 개인 정체성 판단 ■ 적절한 감독 없는 전투 피해 추정 ■ 이민/망명 최종 결정 ■ AI 단독 기반 정보분석 보고 ■ 핵무기 관련 결정에서 인간 배제
고영향 AI 사용 *특별 안전장치 필요	■ 생체인식 기반 실시간 개인 추적 ■ 테러리스트/안보위협 인물 분류 ■ 이민/망명 관련 결정 ■ 위험물질/시스템 관리 ■ 자동 코드 작성/수정 ■ 단독 정보분석 생산연방 인사 관련 AI
연방 인사에 영향을 미치는 AI 사용	■ 채용 결정/임금 책정 ■ 승진/강등/해고 결정 ■ 직무성과/건강상태 평가추가 규제
추가 사용 제한	■ 부서장 재량으로 추가 제한 가능 ■ 공개/비공개 목록 관리 ■ APNSA 보고 의무
Ⅱ. 고영향 AI 및 연방인사 관련 AI 사용에 대한 필수 위험 관리 프레임워크	
위험 및 영향 평가와 효과적인 인간 감독 보장을 위한 기본 위험 관리 요구사항 (180일 내 시행)	■ AI 리스크/영향 평가 - 목적 및 기대효과 : 정량적/정성적 분석 포함, 대안과 비교한 우수성 입증 - 잠재적 위험 평가 : 실패 가능성 문성화, 위험 대비 이점 분석 - 데이터 품질/적합성 : 출처 및 신뢰성 검증, 활용 목적 적합성 확인, 실제 상황 대응력 검증 ■ 운영 및 감독 - 실제 환경 테스트 필수, 독립적 평가 시행, 차별/편향 요소 식별 및 제거, 과도한 AI 의존 방지, 운영자 교육 및 평가, 인간 감독 체계 구축, 문제 보고 시스템 운영 필요 ■ 지속적 관리 - 정기적 모니터링/테스트, 주기적 인적 검토, 새로운 위험 요소 대응, 고위험 사안의 내부 보고체계 유지
연방 인사에 영향을 미치는 AI에 대한 추가 절차적 안전장치 마련	■ 직원 의견 수렴 및 반영, 개인 동의 획득, 불리한 결정 시 통보 의무, 이의제기 절차 보장

Ⅰ. AI 사용 규제	
예외 승인	■ 기본 규정 : 최대 1년 유효 (갱신 가능), AI 책임자 전담 승인, 권한 위임 불가 ■ 관리 의무 : 중앙 집중식 추적 관리, 3일 내 상부 보고, 정기적 재평가 실시, 연간 공개 보고서 발행 ■ 예외 조건 : 프라이버시/시민권/안전 위험 증가 시, 중요 기관 운영 저해 우려 시, 국가안보에 심각한 위험 시
Ⅲ. AI 사용의 목록화 및 모니터링	
목록관리	■ 대상 기관들은 예외 승인 하의 운영을 포함하여 고영향 AI 사용 사례의 연간 목록을 작성하여 APNSA에 보고 ■ 포함 내용 : 각 AI 사용 사례에 대한 설명, 목적과 기대 효과, 해당 사용이 초래하는 위험과 기관의 위험 관리 방법 ※ 부서장들은 목록의 범위, 시기, 메커니즘 및 내용에 대한 상세 지침을 주기적으로 발행하고 업데이트
데이터 관리	■ AI 프레임워크 발행 후 270일 이내에 부서장들은 AI 시스템의 고유한 특성을 고려해 데이터 관리 정책과 절차를 수립하거나 업데이트 필요 ■ 포함내용 : AI 훈련 데이터의 견고성, 대표성, 편향성 평가, 훈련 데이터, 프롬프트에 대한 모범 사례와 표준화, AI 모델의 문서화, 관리, 보존, 중요 결정에 관한 AI 사용 지침, 시민의 자유, 프라이버시, 인권 보호 지침, AI 평가 및 감사 기준, 사이버보안 위험 완화를 위한 지침
감독 및 투명성	■ AI 프레임워크 발행 후 60일 이내 시행해야 하며 부서장들은 이하를 위해 충분한 정보, 전문성, 교육, 자금을 보장해야함
최고 AI 책임자 임명	■ 최고 AI 책임자의 역할 - AI 관련 고위 자문역 수행 - AI NSM 준수를 위한 거버넌스 구축 - AI 활동 모니터링 - 자원 요구사항 검토 - 기관의 AI 정책 조율 - 표준화 활동 지원 - 공정성과 포용성 증진
AI 거버넌스 위원회 설립	■ 최고 AI 책임자가 의장 ■ IT, 사이버보안, 데이터, 프라이버시 등 관련 고위 관리 포함정기적 성과 평가
감독 책임자 지정	■ 프라이버시, 시민의 자유, 투명성, 안전성 감독 ■ AI 활동 문서화 ■ 오용 사례 모니터링 ■ 이해관계자 피드백 수렴 ■ 연간 보고서 작성 (가능한 한 비기밀로)

〈참고 자료〉

1. Biden Administration Outlines Government 'Guardrails' for A.I. Tools,
   The newyork times(2024.10.24.),https://www.nytimes.com/2024/10/24/us/politics/biden-government-guidelines-ai.html
2. Biden Administration Outlines Government 'Guardrails' for A.I. Tools, The New York Times(2024.10.24.),
   https://www.nytimes.com/2024/10/24/us/politics/biden-government-guidelines-ai.html
3. China's AI industry could see US$1.4 trillion in investment in 6 years, executive says, SCMP(2024.9.9.),
   https://www.scmp.com/tech/big-tech/article/3277743/chinas-ai-industry-could-see-us14-trillion-investment-6-years-executive-says
4. Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence, The white house(2024.10.24.),
   https://www.whitehouse.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/
5. Remarks by APNSA Jake Sullivan on AI and National Security, The white house(2024.10.24.),
   https://www.whitehouse.gov/briefing-room/speeches-remarks/2024/10/24/remarks-by-apnsa-jake-sullivan-on-ai-and-national-security/
6. The Biden Administration's National Security Memorandum on AI Explained, CSIS(2024.10.25.).
   https://www.csis.org/analysis/biden-administrations-national-security-memorandum-ai-explained
7. The National Security Memorandum on Artificial Intelligence — CSET Experts React, CSET(2024.10.24.),
   https://cset.georgetown.edu/article/the-national-security-memorandum-on-artificial-intelligence-cset-experts-react/
8. The U.S. National Security Memorandum on AI: Leading Experts Weigh In, Just Security(2024.10.25.),
   https://www.justsecurity.org/104242/memorandum-ai-national-security/
9. Treasury Issues Regulations to Implement Executive Order Addressing U.S. Investments in Certain National Security Technologies and Products in Countries of Concern, U.S. Department of the Treasury(2024.10.28.),
   https://home.treasury.gov/news/press-releases/jy2687
10. U.S. and China Talk AI Safety in Geneva Meeting, Inc(2024.5.13.),
    https://www.inc.com/reuters/us-china-talk-ai-safety-in-geneva-meeting.html
11. US issues AI national security memo to avoid 'strategic surprise' by China and cut risks, SCMP(2024.10.25.),
    https://www.scmp.com/news/china/diplomacy/article/3283702/us-unveil-ai-national-security-memo-avoid-chinas-strategic-surprise-and-cut-risks
12. White House orders Pentagon and intel agencies to increase use of AI, The washington post(2024.10.24.),
    https://www.washingtonpost.com/technology/2024/10/24/white-house-ai-nation-security-memo/
13. "美中 AI 역량, 韓日 등 '2위 그룹'과 격차↑…中 논문 美 추월", 한국경제(2024.7.7.), https://www.hankyung.com/article/202407076945Y
14. 미국, 對중국 AI 반도체 기술 통제 등 추가 규제 도입 검토, KOTRA 경제통상리포트 US(2024.6.21.), 워싱턴무역관 정연호
15. 미중, 'AI 위험 공동 관리' 논의 착수…제네바서 첫 고위급 회담, 연합뉴스(2024.5.15.),
    https://www.yna.co.kr/view/AKR20240515031300009
16. 해외 AI 안전연구소 추진 현황과 시사점, SPRI ISSUE REPORT(2024.7.23.), 유재홍 외

THE AI REPORT 2024

NIA 한국지능정보사회진흥원