

## 백악관 인공지능(AI) 개발 및 활용에 관한 행정명령과 관련 동향

- ◆ 백악관, '안전하고 신뢰할 수 있는 AI의 개발 및 활용에 관한 행정명령' 발표
- ◆ AI 모델 개발 시 정부에 위험 통지 의무화, 안전·보안·신뢰의 AI 혁신 선도 노력
- ◆ 美 의회의 초당적 입법 지원과 G7 국가들 간 협업 및 동참 노력 중요

### I. 개요

#### □ 백악관 인공지능(AI) 행정명령 개요

- '안전하고 신뢰할 수 있는 AI의 개발 및 활용에 관한 행정명령' 발표(10.30)
  - \* Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence
  - 이번 명령은 국방물자생산법(DPA)를 기반으로, AI 모델이 국가 안보나 경제, 공중 보건 관련 위험을 초래할 경우 연방 정부 통지를 의무화
  - 위험한 생물학 무기 개발 방지를 위한 스크리닝 표준 개발과 오남용 및 사기 방지를 위한 콘텐츠 워터마크\* 등 인증 방침 개발 지시
  - \* 알파벳, 메타, 마이크로소프트, 오픈AI 등 주요 AI 기업 자발적 약속('23.7월)
  - AI 보안 관련 추가행동 촉진, 개인정보 보호 기술 개발 가속화 지원, 데이터 개인정보 보호법 통과 촉구, AI 기술 혁신 지원 방안 등 포함되었으며, 같은 날 미국·일본 정부는 주요 7개국(G7) AI 개발자 국제 지침 및 규범 합의 발표

<'히로시마 AI 프로세스' 국제 지침 및 규범 합의('23.10.30)>

- 기존 OECD AI 원칙을 기반으로 11개 항목의 개발자 국제 지침 및 규범 합의 발표
- 생성형 AI 챗GPT 등 첨단 AI의 기회와 변혁 가능성 강조, 위험 관리 필요성 공감
- 국제 행동 규범 준수 요구, 연말까지 히로시마 AI 프로세스 개선 작업 지시
- 콘텐츠 식별 인증 메커니즘 도입, 국제 AI 기술 표준 개발 추진, 데이터 및 지식 재산권 보호 등 안건 포함

## II. 주요 내용

### □ 백악관 AI 행정명령(10.30) 및 예산관리국(OMB) 시행 지침(11.1) 잇따라 발표

- 백악관은 보도자료에서 “미국 정부는 AI의 안전, 보안, 신뢰를 중점으로 AI의 이점을 활용하고 위험을 완화하기 위해 세계 정부 중 가장 강력한 조치를 취하고 있다”며 안전하고 책임 있는 AI 혁신 선도 노력 강조

#### <백악관 AI 행정명령 주요 내용>

- 
- **AI 안전 및 보안 표준 개발(New Standards for AI Safety and Security)**
    - 위험 요소에 대해 사전 연방정부 통보 의무화, AI 시스템 활용을 위해 안전 및 보안 표준 도구 개발, 생물학적 합성 스크리닝 개발 지시
    - AI 생성 콘텐츠 표준 및 모범사례 확립 노력, AI 콘텐츠 인증 및 워터마크에 대한 지침 개발, 사이버 보안 프로그램 구축, 국가안보각서 개발 등
  - **미국인의 개인정보보호(Protecting Americans' Privacy)**
    - 미 의회에 개인정보 보호 법안 통과 촉구, 암호화 도구 및 보호기술 강화
  - **형평성 및 시민권 증진(Advancing Equity and Civil Rights)**
    - AI 권리장전(Blueprint for an AI Bill of Rights)을 기반으로 AI 알고리즘의 차별 방지 대책 마련, AI 민권 침해 조사 및 공정성 보장 노력
  - **AI 보급 지원(Standing Up for Consumers, Patients, and Students)**
    - 의료분야 유해성 보고 체계 수립, AI 지원교육 도구 배포 지원
  - **근로자 지원(Supporting Workers)**
    - 생산성 향상 및 위험완화, 단체교섭능력 지원, 관련 지침 개발 노력
  - **혁신 및 경쟁 촉진(Promoting Innovation and Competition)**
    - 의료 및 기후변화 관련 AI 연구에 보조금 확대 및 연구 촉진, AI 혁신 상용화 장려, 관련 숙련 이민자 및 비이민자 지원 프로그램 개선 등
  - **국제협력 및 리더십(Advancing American Leadership Abroad)**
    - 국제 프레임워크 구축 및 현대화, 표준 개발 구현 주도 및 파트너십 강화
  - **정부기관의 역할(Ensuring Responsible and Effective Government Use of AI)**
    - 명확한 표준 마련으로 AI 조달 개선 및 AI 보급 강화, 간소화 된 조달 계약을 기반으로 AI 제품 및 서비스 지원, AI 전문가 채용 가속화 등
- 

[자료: 백악관 팩트시트]

- OMB 시행지침에서 연방 정부는 인공지능의 안전하고 신뢰할 수 있는 개발 및 사용에 관한 모범을 보이고 기술의 책임 있는 사용에 대한 모델을 제공
  - 세부 목표로 △AI 거버넌스 구조 확립, △책임 있는 AI 혁신 발전, △투명성 제고, △연방 직원 보호, △정부의 AI 사용으로 인한 위험 관리 제시
  - 에너지부는 AI와 화학적, 생물학적, 방사선 위협의 교차점에서 위험을 감소시키고, AI 안전 및 보안을 위한 지침 개발, 컴퓨팅 클러스터에 대한 기술 조건 설정, 민감 데이터 관리, 슈퍼컴퓨터 국가 AI 연구자원 지원 등 부서의 AI 비전 제시
- 미 상무부, NIST 산하 미국인공지능안전연구소(USAISI) 설립 발표(11.1)
  - \* U.S. Artificial Intelligence Safety Institute
  - 백악관 행정명령에 따라 국립표준기술연구소(NIST) 산하 USAISI를 신설하고, AI 모델의 안전, 보안 및 평가를 위한 표준 개발을 촉진
  - 콘텐츠 인증을 위한 표준 개발을 담당하고 연구자로 하여금 새로운 AI 위험을 평가하고 알려진 영향을 해결할 수 있는 테스트 환경을 제공할 전망
  - 영국의 AI 안전 연구소 등 동맹국 및 파트너 국가의 유사한 연구소와 협력하여 관련 분야의 작업을 조정하고 개발하는 업무를 담당

<영국 인공지능 안전 서밋(AI Safety Summit)>

- 
- 영국의 주도로 인공지능의 안전한 사용을 모색하고 인공지능으로 인한 위험을 공동으로 관리하기 위해 미국, EU, 중국 등 참석( '23.11.1~2), 각 국 정상 및 기업 대표 참석
  - 블렛츨리 선언(Bletchley Declaration) 서명: 미국과 중국, EU를 포함한 25개 이상 국가 참석, AI기술 감독에 대한 공통 접근 방식 확립에 합의
  - 공동선언문 채택을 기반으로 두 가지 의제 제시
    - : 1) AI의 안전 위험 식별, 2) 국가별 위험 기반 정책 구축 및 민간 기업 협력
  - 향후 6개월 마다 AI 안전 서밋을 개최하고 다음 개최지는 한국 예정
- 

[자료: 영국 정부보도자료, 파이낸셜타임즈 등]

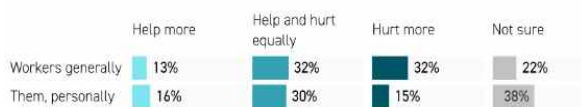
### III. 현지 반응 및 시사점

- 현지 언론, 이번 행정 명령의 주요 초점으로 '국가 안보 도구로서의 AI 개발 (AI as a national security tool)' 주목, 정부 역할 중요성 언급
- 미국 증권거래위원회(SEC) 위원장 게리 겐슬러(Gary Gensler)는 "규제 당국이 AI 위험을 제대로 관리하지 못하면 10년 내 금융 위기를 경험하게 될 것"이라며 정부 주도의 AI 위험 관리 중요성 지목
- 지금까지는 AI 기업 간 우위 선점을 위해 일반적으로 모델 설계를 비공개로 유지해왔으나, 이후 관련 지침을 준수하지 않는 기업의 경우, 미국 정부가 적절한 방식으로 관계 기관 조치가 가능해질 전망
- 인공지능 분야에서 상무부 산하 국립표준기술원(National Institute of Standards and Technology, NIST) 역할 강화 및 AI 안전연구소 설립, AI 시스템 테스트, 표준과 관련된 업계 벤치마크가 강화될 것으로 예상(270일 이내 개발 명시)
- 이번 행정명령의 법적 구속력 및 이행 강화를 위해서는 미 의회의 초당적 입법과 더불어 같은 날 발표된 G7 국가들의 AI 지침(Guiding Principles) 및 행동 강령(Code of Conduct) 간 협업 및 동참 노력\*이 중요해질 전망
- \* 예) 영국의 '프론티어 AI 태스크포스(Frontier AI Taskforce)'과 협업 강화 예상(폴리티코)

<[참고] 미국인의 인공지능(AI) 관련 설문조사('22년 12월)>

#### Nearly one third of Americans believe AI will harm workers more than it will help

How U.S. adults think the use of AI will affect workers and them personally, by percent



Note: Survey conducted in December 2022 among 11,004 U.S. adults. Margin of error is +/- 1 percentage point.

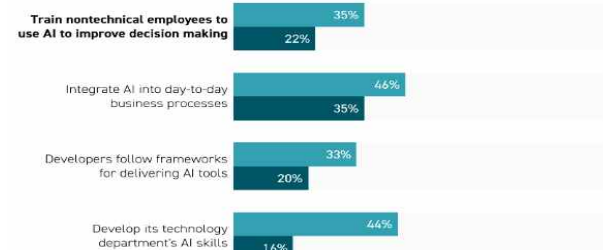
- \* 응답자의 약 3분의 1인 32%는 AI가 근로자에게 긍정적 영향 보단 부정적 영향을 끼칠 것 응답

[자료: Pew Research Center]

#### Before ChatGPT, most employees said their organizations did not offer training in AI

In May and August 2022, percent of employees who said their organization had used AI in the following ways:

■ Respondents whose organizations report at least 20 percent of earnings from AI use  
■ All other respondents



- \* 응답자 대부분 AI 관련 연수를 받아본 적 없음, AI 관련 제도 구현의 느린 속도 지적

[자료: McKinsey & Company]

[자료 : 백악관 보도자료, 현지 언론(폴리티코, WSJ, 인사이드 등) 및 KOIRA 보유자료 종합]