미국 대통령실 관리예산국(OMB)는 2023. 11. 1. <연방기관의 인공지능 사용을 위한 거버넌스, 혁신 및 위험 관리를 개선하기 위한 각서> 초안을 발표하였습니다. OMB AI 각서(규칙)는 국민의 권리와 안전에 영향을 미치는 인공지능을 정의하고 연방정부 기관들이 사용하는 인공지능에 대하여 위험관리를 요구하였습니다. 이 규칙은 민간에서 조달되는 연방정부 AI들에도 적용될 예정입니다.

OMB AI 각서는 2023. 10. 30. 미국 바이든 대통령의 AI 행정명령(Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 안전하고 보안성이 높으며 신뢰할 수 있는 인공지능의 개발 및 사용)에 따라 마련된 것입니다.

OMB는 각서 초안에 대한 사회 각계 의견을 수렴하고 2024. 3. 28. 규칙을 확정하였습니다.
아래 자료는 확정되기 전의 OMB AI규칙의 초안을 번역한 것입니다.

번역: 정보인권연구소 (초벌번역은 기계번역의 도움을 받았습니다. 일부 각주 제외)

공개 검토용 초안 (2023. 11. 1.)

미국 대통령실
관리예산국
Office of Management and Budget (OMB)
워싱턴 **D. C. 20503**

행정부 및 기관장을 위한 제안 각서
Proposed Memorandum for the Heads of Executive Departments and Agencies

발신: 샬란다 D. 영

제목: 기관의 인공지능 사용을 위한 거버넌스, 혁신 및 위험 관리 개선

인공지능(AI)은 우리 시대의 가장 강력한 기술 중 하나이며, 대통령은 AI가 제공하는 기회를 포착하는 동시에 위험을 관리해야 한다는 점을 분명히 했습니다. **2020**년 정부

인공지능법[1], 미국 **AI** 발전법[2],그리고 **2023**년 **10**월 **30**일자 바이든 대통령의 행정명령(인공지능의 안전하고 보안성이 높으며 신뢰할 수 있는 개발 및 사용)에 따라 이 각서는 각 기관에 AI 거버넌스와 혁신을 증진시키는 동시에 AI 사용으로 인한 위험, 특히 공공의 안전과 권리에 영향을 미치는 위험을 관리하도록 지시합니다.

첨부된 연방 관보 공시에 명시된 바와 같이, 관리예산실(OMB)은 이 제안 각서에 대하여 공개적인 의견을 요청합니다.

1. 개요

AI가 연방 정부 전반의 운영과 효율성을 개선하고 있지만, 각 기관은 AI의 사용을 효과적으로 관리해야 합니다. 따라서 이 각서는 대중의 권리와 안전에 영향을 미치는 AI 사용에 대하여 구체적으로 최소한의 위험 관리 관행을 포함하여 AI 거버넌스, 혁신 및 위험 관리에 대한 새로운 기관 요구사항 및 지침을 수립합니다.

AI 거버넌스 강화**.** AI 위험을 관리하고 AI 혁신을 촉진하려면 효과적인 AI 거버넌스가 필요합니다. 바이든 대통령의 2023년 10월 30일 행정명령("AI 행정명령")에 따라 각 기관은 본 각서 발행일로부터 60일 이내에 최고AI책임자(Chief AI Officer, CAIO)를 지정해야 합니다. 이 각서에는 기관 CAIO의 역할, 책임, 연공서열, 직위 및 보고 구조가 설명되어 있습니다. AI는 데이터, 정보 기술(IT), 보안, 개인정보 보호, 시민권 및 시민 자유, 고객 경험, 인력 관리 등 여러 기술 및 정책 분야와 긴밀하게 연결되어 있기 때문에 CAIO는 해당 기관 내 기존 담당 공무원 및 조직과 긴밀히 협력해야 합니다.

책임 있는 **AI** 혁신의 증진**.** 책임 있게 AI를 구현하면 연방 정부 전반의 운영을 개선할 수 있습니다. 각 기관은 생성형 AI를 포함한 AI를 성공적이고 책임 있게 운영에 도입할 수 있는 역량을 키워야 합니다. 이를 위해 이 각서는 최고재무책임자(CFO)법[3]에 명시된 각 기관이 전사적 AI 전략을 개발할 것을 요구합니다. 또한 이 각서는 IT 인프라, 데이터, 사이버 보안, 인력 및 생성형 AI의 특정 과제와 관련된 장벽을 비롯하여 기관이 책임 있는 AI 사용에 대한 장벽을 어떻게 줄여야 하는지에 대한 권장 사항도 제공합니다.

AI 사용으로 인한 위험 관리**.** 기관은 AI를 통해 상당한 이점을 누리는 동시에 AI 사용으로 인한 다양한 위험도 관리해야 합니다. 기관은 AI와 관련된 기존 위험 관리 요구사항의 적용을 받으며, 이 각서는 이러한 요구사항을 대신하거나 대체하지 않습니다. 그보다 이 각서는, 기관의 결정과 조치를 알리거나 수행하기 위해 AI에 의존할 때 발생하는 위험에 초점을 맞춘 새로운 요구사항을 창출합니다. 특히 그러한 의존이 대중의 권리와 안전에 영향을 미치는 경우 그렇습니다[4]. 이러한 위험을 해결하기

[1] Pub. L. No. 116-260, div. U, title 1, § 104 (codified at 40 U. S. C. § 11301 note), https://www. congress. gov/116/plaws/publ260/PLAW-116publ260. pdf.

[2] Pub. L. No. 117-263, div. G, title LXXII, subtitle B, §§ 7224(a), 7224(d)(1)(B), and 7225 (codified at 40 U. S. C. 11301 note), https://www. congress. gov/117/plaws/publ263/PLAW-117publ263. pdf.

[3] 31 U. S. C. § 901(b).

[4] A full definition for "risks from the use of AI" is provided in Section 6.

위해 이 각서는 기관이 권리와 안전에 영향을 미치는 AI를 사용할 때 최소한의 업무 관행을 따르도록 요구하고 권리와 안전에 영향을 미치는 것으로 추정되는 특정 AI 범주를 열거합니다. 마지막으로, 이 각서는 연방 조달 맥락에서 AI 위험을 관리하기 위한 일련의 권장 사항도 수립합니다.

2. 범위

기관의 AI 채택은 많은 문제가 따르는데, 일부는 새롭고 AI에 특정한 문제이고 다른 일부는 잘 알려진 문제입니다. 기관은 AI의 모든 측면에 적절한 주의를 기울여야 하지만, 이 각서의 범위는 특히 기관의 AI 사용과 직접적으로 관련된 거버넌스 및 혁신 문제 뿐만 아니라 AI 사용으로 인해 발생하는 위험을 아우릅니다. 이 각서는 연방 정보시스템과 관련한 일반적인 문제 등 AI 사용과 관계없이 나타나는 문제는 다루지 않습니다. 또한, 이 각서는 기업 위험 관리, 정보 자원 관리, 개인정보 보호, 연방 통계 활동, IT 또는 사이버 보안 정책 등, AI에 적용되지만 특별히 초점을 맞추지 않는 다른 보다 일반적인 연방 정책을 대체하지 않습니다. 기관은 AI와 관련된 여타 영역에서 적용 가능한 OMB 정책을 계속 준수해야 하며 모든 해당 공무원과 기관 전체의 규정 준수를 조정해야 합니다. 모든 기관 책임 공무원은 다른 법률 및 정책에 명시된 기존 권한과 책임을 유지합니다.

a. 해당 기관. 특별히 명시된 경우를 제외하고, 본 각서는 44 U. S. C. § 3502(1)에 정의된 모든 기관에 적용됩니다[5]. 관련 섹션에 명시된 바와 같이, 본 각서의 일부 요구사항은 31 U. S. C. § 901(b)에서 정한 CFO법 적용 기관에만 적용됩니다. 그리고 기타 요구사항은 50 U. S. C. § 3003에 정의된 정보 커뮤니티의 구성원에 적용되지 않습니다.

b. 해당 AI. 본 각서는 아래에 자세히 설명된 대로 해당 기관에 의해 또는 해당 기관을 대신하여 개발, 사용 또는 조달되는(developed, used, or procured by or on behalf of covered agencies.) 신규 및 기존 AI에 적용되는 요구사항 및 권장 사항을 제시합니다. 반면, 본 각서의 원칙은 비기관의 AI 사용에 관한 법률이나 정책을 정하기 위해 고안된 기관의 규제 조치를 규율하지 않습니다.

본 각서의 요구사항은 AI를 통합한 정보 시스템 전체가 아니라 AI를 구현하거나 AI에 의존하는 시스템 기능에 적용됩니다. 관련 섹션에 명시된 바와 같이, 본 각서의 일부 요구사항은 AI가 권리나 안전에 영향을 미칠 수 있는 경우 등 기관이 AI를 사용하는 특정 상황에만 적용됩니다.

[5] The term “agency” is defined as “any executive department, military department, Government corporation, Government controlled corporation, or other establishment in the executive branch of the Government (including the Executive Office of the President), or any independent regulatory agency,” but does not include the Government Accountability Office; the Federal Election Commission; the governments of the District of Columbia and of the territories and possessions of the United States, and their various subdivisions; or Government-owned contractor-operated facilities, including laboratories engaged in national defense research and production activities. As a result, agencies defined in 44 U. S. C. § 3502(5) (independent regulatory agencies) that were not covered by Executive Order 13960 of December 8, 2020 are covered by this memorandum.

c. 국가 보안 시스템에 대한 적용 가능성. AI가 국가 안보 시스템의 구성 요소로 사용되는 경우 본 각서는 해당 내용을 다루지 않습니다. [6]

3. 인공지능 거버넌스 강화

각 해당 기관의 장은 AI 혁신을 추구해야 할 책임 뿐 아니라 해당 기관의 AI 사용으로 인한 위험을 적절하게 관리하는 것을 포함하여 해당 기관이 관련 법률 및 정책의 AI 요구사항을 준수하도록 보장할 책임이 있습니다. 각 해당 기관의 장은 또한 본 각서에 명시된 책임을 지원하기 위해 예산 절차를 통해 자원을 제공하거나 요청하는 것을 포함하여, 이러한 책임을 효과적으로 수행하는 데 필요한 재정 자원, 인적 자원, 정보 자원 및 인프라 자원을 검토해야 합니다.

AI 문제에 대한 책임성을 강화하기 위해 기관은 AI 행정 명령 섹션 10. 1(b)에 따라 최고AI책임자(CAIO)를 지정해야 합니다. CAIO는 이 각서를 이행하고 다른 기관의 이행을 조정하는 데 있어 해당 기관의 장을 대신하여 일차적인 책임을 집니다. 이 섹션에서는 CAIO의 역할, 책임, 연공서열, 직위 및 보고 구조를 정의합니다.

a. 조치

i. 최고**AI** 책임자**(CAIO)** 지정**.** 이 각서가 발행된 후 60일 이내에 각 기관의 장은 CAIO를 지정해야 합니다. CAIO가 본 각서에 명시된 책임을 이행할 수 있도록 이미 CAIO를 지정한 기관은 해당 개인에게 추가 권한을 부여해야 할지 아니면 새로운 CAIO를 임명해야 할지 평가해야 합니다. 기관은 OMB의 통합 데이터 수집 절차 또는 OMB가 지정한 후속 절차를 통해 이들 담당자를 OMB에 알려야 하며, 지정된 개인이 변경되면 30일 이내에 OMB를 업데이트해야 합니다.

ii. 기관 **AI** 거버넌스 조직 소집**.** 본 각서 발행일로부터 60일 이내에 각 CFO법 적용 기관은 AI 행정 명령 섹션 10. 1(b) 및 이 각서 섹션 3(c)의 세부 지침에 따라 AI 문제를 조정하고 관리하기 위해 관련 고위 관리들을 소집해야 합니다.

iii. 규정 준수 계획**.** 정부 내 AI법(AI in Government Act) 104(c)-(d)항에 따라, 본 각서 발동 또는 본 각서 업데이트 후 180일 이내에, 그리고 그 후 2036년까지 2년마다, 각 기관은 본 각서와의 일관성을 달성하기 위한 계획 또는 해당 기관이 해당 AI를 사용하지 않으며 사용할 것으로 예상되지 않는다는 서면 결정을 OMB에 제출하고 해당 기관의 웹사이트에 공개적으로 게시해야 합니다. 기관은 또한 이 각서와의 일관성을 보장하기

[6] AI innovation and risk for national security systems must be managed appropriately, but these systems are governed through other policy. For example, Section 4. 8 of the AI Executive Order requires the development of a National Security Memorandum to govern the use of AI as a component of a National Security System, and agencies have existing guidelines in place such as the Department of Defense's (DoD) Responsible Artificial Intelligence Strategy and Implementation Pathway and the Office of the Director of National Intelligence's Principles of Artificial Intelligence Ethics for the Intelligence Community, as well as policies governing specific high-risk national security applications of AI, such as DoD Directive 3000. 09, Autonomy in Weapon Systems.

위해 기존 내부 AI 원칙 및 지침을 업데이트할 계획을 포함해야 합니다[7]. OMB는 이러한 준수 계획에 대한 전체적인 템플릿을 제공할 것입니다.

iv. AI 사용 사례 인벤토리**.** 미국 AI 발전법(Advancing American AI Act) 섹션 7225에 따라 해당 법과 AI 행정 명령 섹션 10. 1(e)에서 제외되는 경우 각 기관(국방부 및 정보 커뮤니티 제외)은 매년 AI 사용 사례 목록(인벤토리)을 OMB에 제출한 후 해당 기관의 웹사이트에 공개 버전을 게시하여야 합니다[8]. OMB는 통합 데이터 수집 절차 또는 OMB가 지정한 후속 절차를 통해 인벤토리에 대한 자세한 지침을 발행할 것입니다. 2024년 사용 사례 인벤토리부터 시작해서 해당하는 기관인 경우 안전에 영향을 미치고 권리에 영향을 미치는 AI를 사용하는 자신들의 방법, 이러한 사용으로 인해 발생하는 위험(형평성에 대한 위험 포함), 해당 위험을 관리하는 자신들의 방법, 이 각서 섹션 5에서 부여되는 관련 추가사항과 면제와 관련한 내용에 대한 추가 세부 정보를 파악하고 보고해야 합니다.

v. 인벤토리 대상이 아닌 **AI** 사용 사례 보고**.** 일부 AI 사용 사례는 미국 AI 발전법의 인벤토리 요구사항이 면제됩니다. 이러한 사용 사례 중 국방부 내부 사례는 국가 보안 시스템의 구성 요소로 사용되는 AI와 관련되지 않는 한 이 각서의 범위 내에 있습니다. 국방부는 범위 내 AI 사용 사례에 대한 집계 지표, 권리와 안전에 영향을 미치는 사례 수, 본 각서 섹션 5 (c) 의 관행 준수 및 본 각서 섹션 5에 따라 부여된 면제를 포함하여 범위 내 AI 사용 사례에 대한 정보를 OMB에 매년 제공해야 합니다. OMB는 통합 데이터 수집 절차 또는 OMB가 지정한 후속 절차를 통해 이 보고에 대한 자세한 지침을 발행할 것입니다.

b. CAIO의 역할**,** 책임**,** 연공서열**,** 직위 및 보고 구조

AI 행정명령 섹션 10. 1(b)(ii)에 따라 본 각서는 기관 CAIO의 역할, 책임, 연공서열, 직위 및 보고 구조를 다음과 같이 정의합니다.

i. 역할**.** CAIO는 이 섹션에 설명된 책임을 수행하는 데 필요한 기술, 지식, 교육훈련이력 및 전문성을 갖추고 있어야 합니다. CFO법 적용 기관에서 CAIO의 주요 역할은 해당 기관의 AI 사용에 대한 조정, 혁신 및 위험 관리에 있어야 합니다. 기관은 최고기술책임자, 최고데이터 책임자, 관련되어 있거나 보조적인 권한과 책임을 보유한 유사 업무 공무원 등 기존의 공무원이 AI에 대한 상당한 전문성을 갖추고 이 섹션의 다른 요구사항을 충족하는 경우 CAIO로 지정할 수 있습니다.

ii. 책임**.** AI 행정 명령은 해당 기관에서 다른 담당 공무원과 협력하여 기관의 AI 사용을 조정하고, AI 혁신을 촉진하고, AI 사용으로 인한 위험을 관리하고, 행정명령

[7] Given the importance of context-specific guidance on AI, agencies are encouraged to continue implementing their agency's AI principles and guidelines, so long as they do not conflict with the guidance in this memorandum.

[8] Agencies must only publicly report use cases to the extent practicable and consistent with applicable law and governmentwide guidance, including those concerning the protection of privacy and of sensitive law enforcement, national security, and other protected information.

13960호 섹션 8(c)[9] 및 행정명령 14091의 섹션 4(b)[10]에 정의된 기관 책임을 수행하는 일차적인 책임을 CAIO에 부여합니다. 또한 CAIO는 다른 담당 공무원 및 적절한 이해관계자와 협력하여 다음 사항을 책임집니다.

[기관의 AI 사용을 조정함]

A. 기관장 및 기타 기관 고위진에 대하여 그리고 기관의 고위 의사결정 회의에서 AI에 대한 수석 자문 역할을 수행합니다.

B. 연간 AI 사용 사례 인벤토리 작성 및 유지를 포함하여 기관 AI 활동에 대한 인식을 유지합니다.

C. 본 각서 섹션 3(a)(iii)에 자세히 설명된 대로 본 각서의 준수 계획과 본 각서의 섹션 4(a)에 자세히 설명된 기관 AI 전략을 개발합니다.

D. 기관의 임무에 AI를 적용하고 위험을 적절하게 관리하는 데 필요한 자원할당 요구사항 및 인력 스킬셋(workforce skillsets)에 대해 기관 CFO 및 최고인적자본 책임자(CHCO)에 조언합니다.

E. AI 행정 명령 섹션 10. 1(a)에 설명된 위원회에서 기관을 대표하는 것을 포함하여 기관 AI 활동과 관련된 특정 기관 간 조정 기구에 기관이 참여하는 것을 지원합니다.

F. 기관의 AI 표준 설정 기구 참여를 지원하고 조정하며, 해당되는 경우 기관이 OMB 회람문서 No. A-119에 따라 적절하고 일관성 있게 자발적인 AI 합의 표준을 채택하도록 독려합니다[11].

[AI 혁신을 촉진함]

G. 기관에서 협력하여 기관의 임무 수행을 개선하고 형평성을 증진할 수 있는 AI의 적절한 사용 방안을 파악하고 우선순위를 지정합니다.

H. AI 지원 기업 인프라, 인력 개발 조치, 정책 및 기타 AI 혁신 자원의 증진을 포함하여 기관 내 AI의 책임 있는 사용에 대한 장벽을 파악하고 제거합니다.

I. 기관의 사명에서 AI가 가진 기회와 이점을 기관 내적으로 그리고 공공적으로 옹호합니다.

---

[9] Executive Order 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, https://www. govinfo. gov/content/pkg/FR-2020-12-08/pdf/2020-27065. pdf.

[10] Executive Order 14091, Further Advancing Racial Equity and Support for Underserved Communities Through the Federal Government, https://www. govinfo. gov/content/pkg/FR-2023-02-22/pdf/2023-03779. pdf.

[11] OMB Circular A-119, Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities (Feb. 10, 1998), https://www. whitehouse. gov/wpcontent/uploads/2017/11/Circular-119-1. pdf.

[AI 사용으로 인한 위험을 관리함]

J. 특히 안전에 영향을 미치고 권리에 영향을 미치는 AI의 경우 기관이 AI 사용으로 인한 위험을 파악하고 관리하는 것을 지원하는 기관 프로그램을 관리합니다.

K. 관련된 기관 고위 공무원과 협력하여 AI 애플리케이션의 지속적인 성능과 의도한 목표 달성 여부를 측정, 모니터링 및 평가하는 절차를 수립하거나 업데이트합니다.

L. 본 각서 및 관련 법률 및 정책에 명시된 사항을 포함하여 AI 사용으로 인한 위험을 관리하기 위한 요구사항을 기관이 준수하는지 감독합니다.

M. 본 각서에 대한 준수를 보장하기 위해 필요한 경우 기관 AI 애플리케이션의 위험 평가를 수행합니다.

N. 필요한 경우 AI가 안전에 영향을 미치거나 권리에 영향을 미치는 것으로 추정되는 목적이 무엇인지 기관 맞춤 목록 개발을 감독합니다[12].

O. 각서 섹션 5에 자세히 설명된 절차에 따라 해당 요소에서 AI 애플리케이션을 개별적으로 면제합니다.

P. 관련 기관 공무원(예: 허가, 조달, 법률, 인적 자원 및 감독 공무원)과 협력하여, AI 사용 위험에 기반하여 관련 기관 공무원들의 운영 권한 평가 업무를 보조하는 등, 기관이 본 각서를 준수하지 않는 AI를 사용하지 않도록 보장합니다.

iii. 연공 서열**.** CFO법 적용 기관의 경우 CAIO는 고위 집행부, 과학 및 전문직, 고위 경영진 수준 또는 이와 동등한 직위여야 합니다. 다른 기관에서는 CAIO가 최소한 GS-15 또는 이에 상응하는 자격을 갖추어야 합니다.

iv. 직위 및 보고 구조**.** CAIO는 이 섹션의 책임을 수행하는 데 필요한 권한을 가져야 하며, 부기관장 또는 이와 동등한 사람을 포함하여 여타의 기관 경영진과 정기적으로 협력할 수 있을 만큼 높은 위치에 있어야 합니다. 또한 CAIO는 해당 기관의 다른 담당 공무원과 협력하여 기관의 AI 사용이 관련 법률 및 정부 차원 지침을 준수하고 적절하도록 보장해야 합니다.

c. 기관 내부 **AI** 조정

기관은 AI 문제가 기관의 고위 경영진으로부터 적절한 관심을 받도록 해야 합니다. AI 행정명령 섹션 10. 1(b)에 따라 기관은 AI 거버넌스 기구 소집 등을 통해 AI 채택 및 위험 관리 부문을 담당하는 공무원 간 내부 조정을 위한 적절한 조치를 취해야 합니다.

[12] See Section 5(b) of this memorandum for the OMB-defined lists to which agency-specific lists would add. Any agency-specific lists will be governed by the same processes defined in Section 5(b) for the OMB-defined lists.

마찬가지로 CAIO는 기관 위험 관리 전략 개발을 포함하여[13], 적절한 시기에 기관 전반의 위험 관리 기구 및 절차[14]에 참여해야 합니다. 기관의 AI 조정 메커니즘은 기관이 현재 AI를 사용하는 정도, AI가 기관의 임무 수행을 개선할 수 있는 정도, 기관의 현재 및 AI의 잠재적인 용도 등에 기반하여 기관의 요구에 맞추어져야 합니다.

CFO법 적용 기관은 AI 사용에 대한 장벽을 제거하고 관련 위험을 관리하는 등 기관의 AI 사용을 관리하기 위하여, 관련 고위 공무원을 분기별로 소집하는 AI 거버넌스 위원회를 설립하여야 합니다. 해당 기관이 현재 다음 두 가지를 모두 충족하거나 충족하려고 하는 한, 이러한 요구사항을 충족하는 기존 거버넌스 기관에 의존하는 것이 허용됩니다[15].

i. 기관 AI 거버넌스 위원회는 해당 기관의 부기관장 또는 이와 동등한 사람이 의장을 맡고 기관 CAIO가 부의장을 맡아야 하며, 이러한 책임을 다른 공무원에게 할당해서는 안 됩니다. CAIO는 이 위원회를 통해 각자의 부기관장이 기관 전반적으로 AI 활동을 조정하고 AI 행정 명령의 관련 섹션을 구현하는 것을 지원할 것입니다.

ii. 기관 AI 거버넌스 위원회는 최소한 IT, 사이버 보안, 데이터, 인적 자원, 조달, 예산, 기관 관리, 고객 경험, 성과 평가, 통계, 위험 관리, 형평성, 개인정보 보호, 시민권 및 시민 자유를 비롯하여 AI의 채택 및 위험 관리에서 주요한 구현을 책임지는 기관 고위직 공무원의 적절한 대표자와, 기관의 프로그램 사무실 내에서 AI 구현을 담당하는 공무원을 포함해야 합니다. 기관은 또한 각자의 일반 감찰관실의 대표자를 포함하는 것을 고려해야 합니다.

4. 책임 있는 인공지능**(Responsible Artificial Intelligence)** 혁신의 증진

AI가 책임 있게 구현된다면 연방 정부 전반에 걸쳐 운영을 개선하고 효율성을 높일 수 있습니다. 기관은 공공에 혜택이 되고 임무 수행 효율성을 높이는 방식으로 AI를 사용하는 능력을 향상시키되, 다만 AI의 한계와 AI가 특정 작업에 적합하지 않은 경우를 인식해야 합니다. 이를 달성하기 위해 기관은 책임 있는 AI 혁신을 지원할 수 있는 기관 내부 역량을 구축하고 AI 조달을 개선하기 위한 조치를 취해야 합니다.

a. AI 전략

[13] See OMB Circular No. A-130, Managing Information as a Strategic Resource, Appx. I, sec. 5(b) (July 28, 2016), https://www. whitehouse. gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised. pdf.

[14] See, e. g. , OMB Circular No. A-123, Management's Responsibility for Enterprise Risk Management and Internal Control (July 15, 2016), https://www. whitehouse. gov/wpcontent/uploads/legacy_drupal_files/omb/memoranda/2016/m-16-17. pdf.

[15] An example of a qualifying body includes agency Data Governance Bodies, established by OMB Memorandum M-19-23, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, https://www. whitehouse. gov/wp-content/uploads/2019/07/m-1923. pdf.

본 각서 발행 후 365일 이내에 각 CFO법 적용 기관은 책임 있는 AI 사용을 가로막는 장벽을 파악 및 제거하고, 전사적으로 AI 성숙도를 향상시키기 위하여 다음을 포함한 전략을 개발하여 해당 기관의 웹 사이트에 공개해야 합니다.

i. 기관의 현재 및 계획 중인 주요 AI 사용 사례[16]

ii. AI 행정 명령 섹션 10. 1(c)에 따라 수립된 방법에 따라 기관의 AI 성숙도 및 기관의 AI 성숙도 목표에 대한 현재의 평가

iii. CAIO, AI 거버넌스 위원회, AI 사용 사례 인벤토리 개선 등을 통해 기관의 AI 사용을 효과적으로 관리하기 위한 계획

iv. AI를 구축, 테스트 및 유지 관리하는 데 필요한 데이터, 컴퓨팅, 개발, 테스트, 사이버 보안 규정 준수, 배치 및 지속적인 모니터링 인프라에 있어 성숙한 AI 지원 인프라를 포함하여 충분한 기관 AI 혁신 역량을 개발하기 위한 계획

v. AI 사용으로 인한 위험을 관리하기에 충분한 기관의 역량 구축 계획

vi. 기관의 AI 인력 정원 및 예상되는 AI 인력 수요에 대한 현재의 평가, 그리고 이러한 수요를 충족하기 위하여 AI 실무자를 모집, 채용, 교육훈련, 유지, 권한부여를 실시하고 AI에 관여하는 비실무자가 AI 활용 능력(리터러시)을 달성하기 위한 계획

vii. 향후 AI 투자에 있어 구체적이고 우선순위가 높은 분야와 계획

b. 인공지능의 책임 있는 사용을 가로막는 장벽 제거

혁신을 수용하려면 AI 사용에 불필요하고 도움이 되지 않는 장벽을 제거하는 동시에 책임 있는 사용을 보장하는 가드레일을 유지하고 강화해야 합니다. 기관은 AI를 개발하고 배치하는 사람들이 유연성을 갖추는 한편 제한적인 자원과 전문성으로 AI 혁신과 위험 관리에 집중할 수 없는 방해요인에 직면하지 않도록 내부 환경을 조성해야 합니다. 기관은 다음 권장 사항에 특별한 주의를 기울여 이러한 장벽을 제거하기 위한 조치를 취해야 합니다.

i. IT 인프라**.** 기관은 AI 프로젝트가 필요한 경우 AI 학습 및 추론에 특화된 고성능 컴퓨팅 인프라를 포함하여 적절한 IT 인프라에 액세스할 수 있도록 보장해야 합니다. 또한 기관은 AI 개발자가 소프트웨어 도구, 오픈 소스 라이브러리에 적절하게 액세스하도록 보장하고 AI 애플리케이션을 신속하게 개발, 테스트 및 유지 관리하는 데 필요한 역량을 배치 및 모니터링해야 합니다.

ii. 데이터**.** 기관은 AI 학습, 테스트 및 운영에 사용할 기관의 데이터 세트를 충분히 큐레이션할 수 있는 적절한 인프라와 역량을 개발해야 합니다. 여기에는 기관 내부 데이터에 대한 적절한 접근성을 최대화하고 해당 데이터를 내부에 공유할 수 있는

[16] Consistent with Sections 7225(d) and 7228 of the Advancing American AI Act, this requirement applies to CFO Act agencies except for the Department of Defense and the Intelligence Community, as defined in 5 U. S. C. § 3003(4).

기관의 역량이 포함됩니다. 또한 기관은 공개적으로 접근할 수 있는 데이터 세트의 유용성을 탐구하고 그 사용이 적절하고 이 각서에 설명된 데이터 관행에 부합하다면, AI 애플리케이션을 개발, 테스트 및 유지 관리하는 데 도움이 될 수 있도록 그 사용을 권장해야 합니다. 이러한 활동은 특히 데이터 큐레이션, 라벨링, 책무와 관련된 건전한 데이터 거버넌스 및 관리 관행을 가능하게 하는 자원 할당을 통해 뒷받침되어야 합니다.

iii. 사이버 보안**.** 기관은 필요에 따라 사이버 보안 인증 프로세스를 업데이트하여 AI의 지속적인 인증 사용을 증진하는 등 AI 애플리케이션에 대한 요구사항을 더 잘 처리해야 합니다. AI 행정 명령 섹션 10. 1(f)에 따라, 기관의 인증 담당자는 운영 허가 절차 및 기타 해당되는 모든 배포 또는 감독 절차에서 생성형 AI 및 기타 주요 신기술의 우선 순위를 지정해야 합니다.

iv. 인력**.** AI 행정 명령 섹션 5. 1 및 10. 2에 따라 기관은 AI 인재의 공백을 메우고 이용 가능한 특별 채용 및 고용 권한을 최대한 활용하여 다양한 관점과 경험을 가진 개인의 지원을 장려하고 해설 직책 및 기술 기반 평가 등 AI 직책에 대한 채용 모범 사례를 활용하여야 합니다. 기관은 AI에 대한 인력 수요를 파악하고 충원할 때 성공적이고 책임 있는 AI를 달성하기 위하여 AI에 대한 기여하는 바와 그 역량이 중요한 데이터 과학자 및 엔지니어와 같은 기술적 직책과 디자이너, 행동 과학자, 계약 담당자, 관리자, 변호사와 같은 비기술적 직책을 모두 포함해야 합니다. 기관은 내부적으로 AI 인재를 개발하기 위한 자원할당과 교육을 제공해야 하며, 또한 연방 직원에게 AI 직종으로 진출할 수 있는 경로를 제공하거나 업무에 AI가 적용되어 그 영향을 받게 되는 직원을 지원하는 등 연방 직원을 위한 AI 교육 제공을 확대해야 합니다.

v. 생성형 **AI.** AI 행정 명령 섹션 10. 1(f)에 명시된 지침에 유의하는 것 외에도, 기관은 자기관 임무 수행에 있어 생성형 AI의 잠재적으로 유익한 사용 사례를 평가하고 적절한 안전 조치와 감독 메커니즘을 수립하여 기관에서 사용되는 생성형 AI가 과도한 위험을 초래하지 않도록 해야 합니다.

5. 인공지능 사용으로 인한 위험 관리

기관은 기관의 정보 및 시스템과 관련된 위험을 관리하기 위해 다양한 정책, 절차 및 담당 직원을 두고 있습니다. AI 사용으로 인한 위험, 특히 국민의 권리와 안전에 미치는 위험에 더 잘 대처하기 위해 정보 커뮤니티의 구성원이 아닌 모든 기관은, 권리와 안전에 영향을 미치는 AI로 인한 위험을 관리하기 위하여 아래에 자세히 설명된 최소 관행(minimum practices)을 구현해야 합니다[17].

a. 조치

i. 위험 관리 관행의 구현 및 비준수 **AI**의 종료**.** 2024년 8월 1일까지 기관은 안전에 영향을 미치거나 권리에 영향을 미치는 AI에 대하여 본 각서 섹션 5(c)에 명시된 최소

[17] Although elements of the Intelligence Community are not required to implement these practices, they are encouraged to do so.

관행을 구현해야 하며, 그렇지 않은 경우 해당 섹션의 세부 주의사항에 따라 최소 관행을 준수하지 않는 모든 AI의 사용을 중단해야 합니다.

ii. AI 문서화에 대한 권장 사항**.** 본 각서 발행일로부터 180일 이내에 AI 행정 명령 섹션 10. 1(a)에 명시된 위원회는 OMB 국장에게 연방 AI 계약 규정을 이행하기 위해 선정된 공급업체에게 요구해야 하는 권장 문서 목록을 제공할 것입니다. 해당 위원회는 권장 사항의 일부로서 섹션 5(c)의 최소 위험 관리 관행과, 해당 조치를 완수했음을 입증하기 위해 공급업체에 요구할 수 있는 관련 자료를 고려해야 합니다.

b. 어떤 인공지능이 안전에 영향을 미치거나 권리에 영향을 미치는 것으로 추정되는지를 결정

이 섹션 범위 내에서 섹션 6에 정의된 "안전에 영향을 미치는 AI" 또는 "권리에 영향을 미치는 AI"의 정의와 일치하는 모든 AI는, 적절한 기한까지 섹션 5(c)의 최소 관행을 따라야 합니다. 기관은 개발 중이거나 사용 중인 AI의 각 용도를 검토하여 그것이 안전에 영향을 미치거나 권리에 영향을 미치는 AI의 정의와 일치하는지 여부를 확인해야 합니다.

이 하위 섹션의 범주는 AI가 안전에 영향을 미치거나 권리에 영향을 미치는 것으로 자동으로 추정되는 특정 목적의 하위 집합만을 식별할 뿐이며, 안전에 영향을 미치거나 권리에 영향을 미치는 목적의 AI에 대한 전체 목록을 의미하지는 않습니다. 또한 기관은 기관 내에서 안전이나 권리에 영향을 미치는 것으로 추정되는 특정 목적을 정의하여 섹션 5(c)의 관행을 따르도록 하는 것이 바람직합니다. 기관은 그러한 기관별 목록을 매년 OMB에 보고해야 합니다.

기관이 현재 아래 설명된 목적으로 AI를 사용하고 있거나 사용할 계획인 경우, CAIO는 해당 기관이 지정한 다른 관련 공무원과 협력하여, AI 애플리케이션 또는 구성 요소[18]가 "안전에 영향을 미치는 AI" 또는 "권리에 영향을 미치는 AI"의 정의에 부합하지 않으므로 최소 관행이 적용되지 않는다고 결정(또는 이전 결정을 번복)할 수 있습니다. 기관 CAIO는 문서화된 상황별 및 시스템별 위험 평가를 통해서만 이러한 결정을 내리거나 번복할 수 있습니다. 이러한 결정이나 번복은 30일 이내에 OMB에 보고되어야 합니다.

i. 안전에 영향을 미치는 것으로 추정되는 목적**.** CAIO가 달리 결정하지 않는 한, 본 각서의 범위에 포함되는 AI는 안전에 영향을 미치는 것으로 추정됩니다. 해당 AI가 다음 활동의 결과를 제어하거나 중대한 영향을 미치는 데 사용되는 경우, 안전에 영향을 미치는 AI에 대한 최소 관행을 따라야 합니다.

A. 댐, 응급 서비스, 전력망 또는 에너지 생성이나 이동, 화재 안전 시스템, 식품 안전 메커니즘, 선거 또는 투표 인프라의 무결성, 교통 통제 시스템 및 물리적 교통, 상하수도 시스템, 원자로·원자재·폐기물을 제어하는 기타 시스템의 기능

[18] CAIOs may also make these determinations across groups of closely related AI applications or components, provided that: (1) those systems have undergone a risk assessment that adequately considers the risks from each individual system; and (2) the systems are substantially identical in their risk profiles.

B. 직장, 학교, 주택, 교통, 의료 또는 법 집행 환경 내 로봇 부속물이나 동체의 동작과 같이 인간과 로봇이 협업하는 경우 등에서 이루어지는 물리적인 작동

C. 운동력의 적용, 생물학적 또는 화학적 작용의 전달 또는 잠재적으로 유해한 전자기 자극의 전달

D. 육상, 지하, 해상, 공중 또는 우주 등을 불문하는 차량의 이동

E. 유해 화학물질이나 생물학적 개체 또는 전달경로의 운송, 안전, 설계 또는 개발

F. 산업적 배출 및 환경 영향을 관리하는 절차

G. 산업 폐기물 또는 기타 규제 대상 오염 물질의 운송 및 관리

H. 실패할 경우 안전에 중대한 위험을 초래할 수 있는 산업 장비, 시스템 또는 구조물의 설계, 시공 및 테스트

I. 내부자 위협에 대한 대응

J. 정부 시설에 대한 접근 또는 보안

K. 수출, 투자 또는 운송에 대한 제재, 무역 제한 또는 기타 통제에 따른 집행 조치

ii. 권리에 영향을 미치는 것으로 추정되는 목적**.** CAIO가 달리 결정하지 않는 한, 아래 AI는 권리에 영향을 미치는(그리고 잠재적으로 안전에도 영향을 미치는) 것으로 추정됩니다. 해당 AI가 다음 활동 또는 결정 중 하나의 결과를 제어하거나 중대한 영향을 미치는 데 사용되는 경우, 기관은 권리에 영향을 미치는 AI 및 안전에 영향을 미치는 AI에 대한 최소 관행을 따라야 합니다.

A. 보호받는 표현의 도달 범위를 차단, 삭제, 숨기거나 제한하기로 하는 결정

B. 개인에 대한 법 집행 또는 감시 관련 위험의 평가, 범죄 재범의 예측, 범죄자의 예측, 가해자의 신원 예측, 피해자의 예측, 범죄의 예측, 자동차번호판의 판독, 홍채 인식, 얼굴 인식, 얼굴 스케치, 유전적 얼굴의 재구성, 소셜 미디어의 모니터링, 교도소의 모니터링, 포렌식 분석, 포렌식 유전학, 사이버 침입 행위, 물리적 위치 추적 장치. 선고, 가석방, 감독 석방, 보호관찰, 보석, 재판 전 석방 또는 재판 전 구금과 관련된 결정

C. 이민, 망명 또는 구금 상태에 대한 결정. 미국 또는 미국 영토를 여행할 예정이거나 이미 입국한 개인에 대한 위험 평가의 실시. 생체 인식(예: 얼굴 인식) 또는 기타 수단(예: 소셜 미디어 또는 보호받는 온라인 표현에 대한 모니터링)을 통해 출입국이나 연방 이민 서비스에 대한 이용의 결정. 이민, 망명, 구금 또는 출입국 상황에서 개인에 대한 공식 의사소통의 번역. 이민, 망명 또는 구금 관련 물리적 위치 추적 장치.

D. 인간의 감정, 생각 또는 속임수의 감지 및 측정

E. 교육기관에서 학생의 부정 행위 또는 표절의 탐지, 입학 절차에 영향을 미치는 행위, 온라인 또는 가상 현실에서 학생에 대한 모니터링, 학생의 진도 또는 결과에 대한 예상, 징계 중재 권고, 교육 자원 또는 프로그램에 대한 이용의 결정, 학생 지원에 대한 적격 판단, (온라인 또는 대면으로 이루어지는) 촉진 감시

F. 세입자 심사 또는 관리, 주택 평가, 모기지 인수, 주택 보험의 이용 또는 조건의 결정

G. 채용 심사, 급여 또는 승진, 성과 관리, 고용 또는 해고, 근무 시간 추적, 가상 또는 증강 현실 직장 교육 프로그램, 전자 직장 감시 및 관리 시스템을 포함한 고용 조건의 결정

H. 의료 기기, 의료 진단 도구, 임상 진단 및 치료 판단, 의료상 또는 보험상의 건강 위험 평가, 약물 중독 위험의 평가 및 관련 이용 시스템, 자살 또는 기타 폭력 위험의 평가, 정신 건강 상태의 감지 및 예방, 개입 대상 환자에 대한 표시 시스템, 공공 보험 진료 할당 시스템, 건강 보험 비용 및 보증에 대한 절차

I. 대출 할당 절차, 금융 시스템의 이용 결정, 신용 평가, 재무 감사 대상자의 결정, 위험 평가를 포함한 보험 절차, 이자율의 결정, 제재(예: 임금을 압류하거나 세금 환급을 보류하는 등) 대상 금융 시스템

J. 정부 급여나 서비스에 대한 이용, 적격 또는 취소에 관한 결정. 급여 서비스에 접근하는 데 사용되는 IT 시스템의 이용을 생체 인식 또는 기타 수단(예: 서명 일치)을 통해 허용하거나 거부하는 것. 부정수급 탐지. 정부 급여 관련 벌칙의 부과

K. 아동 급여, 아동 양육권, 부모나 후견인이 아동 양육권을 취득하거나 유지하는 데 적합한지 여부에 관한 권고나 결정

c. 안전에 영향을 미치고 권리에 영향을 미치는 인공지능에 대한 최소 관행

관련 법률 및 정부 차원의 지침에 의해 금지되는 경우를 제외하고, 기관은 2024년 8월 1일까지 안전에 영향을 미치고 권리에 영향을 미치는 AI에 대하여 이 섹션의 최소 관행을 적용해야 하며, 그렇지 않으면 AI가 규정을 준수할 때까지 AI 사용을 중단해야 합니다. 2024년 8월 1일 이전에 기관 CAIO는 해당 기관의 관련 공무원과 협력하여 잠재적으로 규정을 준수하지 않는 AI에 대하여 준수 조치를 실시하여야 합니다. 여기에는 제3자 공급업체에 적절한 조치(예: 업데이트된 문서 또는 테스트 조치 실시)를 취하도록 자발적으로 요청하는 것이 포함될 수 있습니다. 이 요구사항을 준수하려면 관련 기관 담당자는 가능한 한 운영 허가 절차 등 기존 메커니즘을 활용해야 합니다. 기관은 CAIO를 통해 아래 설명된 절차를 활용하여 이 요구사항에 대한 면제를 승인하거나 연장해줄 것을 요청할 수도 있습니다.

기관은 이러한 관행의 이행을 문서화하고, 이에 대하여 연간 AI 사용 사례 인벤토리의 구성 요소, TechStat[19] 프로세스 등 정기적인 책임 검토, 또는 OMB가 결정한 요청에 따라 OMB에 보고할 준비가 되어 있어야 합니다.

이 섹션의 사례는 AI 사용으로 인한 위험을 관리하기 위한 최소 기준을 제시합니다. 기관은 결정된 사용 사례와 관련하여 상황별로 추가적인 위험을 파악하고 적절하게 해결해야 합니다. 이러한 위험 고려 사항에는 안전, 보안, 시민권, 시민 자유, 개인정보 보호, 민주적 가치, 인권, 기회 균등, 노동자 복지에 미치는 잠재적 피해, 중요 자원 및 서비스에 대한 접근성, 시장 경쟁에 미치는 영향 등이 포함될 수 있습니다. 잠재적인 위험 관리의 격차를 메우기 위해 기관은 NIST(미 국가표준기술원) AI 위험 관리 프레임워크[20], AI 권리 장전 청사진[21], 적용 가능한 국제 표준[22], AI 행정명령 섹션 6에 따라 수립된 인력 원칙 등 AI 위험 관리에 대한 추가 모범 사례를 적절하게 홍보하고 통합할 것이 권장됩니다. 또한 기관은 본 각서와 행정명령 13960, 행정명령 14091, 2023년 10월 30일 AI 행정명령의 원칙에 따라 적절하고 일관되게 기관별 자체 관행을 계속 개발할 것이 권장됩니다. 또한 이 섹션의 관행은 법률이나 정부 차원의 정책에 따라 요구되는 기존 요구사항의 해석을 대체, 수정 또는 지시하지 않으며, 기관 책임자는 이 관행의 이행이 적용받는 다른 법률이나 정부 차원의 지침과 충돌하지 않도록 조정해야 합니다.

i. 최소 관행 제외**.** AI가 다음 목적 중 하나 이상의 용도로만 사용되는 경우 기관은 이 섹션에 설명된 최소 관행을 따를 필요가 없습니다.

A. 조달 또는 인수 결정을 내리기 위한 목적으로만 기관 기관 운영에 사용된 바 없는 잠재적 공급업체, 상용 기능 또는 무료 제공 AI 기능에 대한 평가

B. AI 제공업체가 규제 집행, 법 집행 또는 국가 안보 조치의 대상이거나 잠재적인 대상이기 때문에 이루어지는 특정 AI 애플리케이션에 대한 평가[23]

C. 연구 개발[24]

---

[19] Policies & Initiatives: TechStat, U.S. Chief Information Officers Council, https://www.cio.gov/handbook/policies-initiatives/techstat/.

[20] Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Publication AI 100-1, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[21] Blueprint for an AI Bill of Rights, White House Office of Science and Technology Policy, https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

[22] For example, ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management, https://www.iso.org/standard/77304.html.

[23] Agencies are not required to follow these minimum practices when examining AI as the target or potential target of such an action, but they are required to follow these practices when carrying out an enforcement or national security action. For example, when evaluating an AI tool to determine whether it violates the law, agencies need not follow the minimum practices; if agencies were using that same tool to assess a different target, they would have to follow the minimum practices.

[24] AI research and development is not excluded if it is used in agency operations other than for the purposes of research and development, such as to make agency recommendations or decisions about real people.

ii. 최소 관행을 위한 기한 연장**.** 2024년 8월 1일까지 기관은 이 섹션의 최소 요구사항을 해당 일자까지 충족할 수 없는 AI의 특정 사용에 대해 제한적이고 명확한 기간에 한해 연장해줄 것을 OMB에 요청할 수 있습니다. 이 요청에는 해당 기관이 문제의 사용 사례에 대한 규정을 준수할 수 없는 이유와 규정 미준수로 인한 위험을 완화하기 위해 기관이 마련한 관행에 대하여 자세한 근거는 물론, 이 섹션에서 요구하는 최소 관행의 전체 세트를 어떻게 이행할 것인지에 대한 계획을 첨부하여야 합니다.

iii. 최소 관행 면제**.** 기관 CAIO는 다른 관련 공무원과 협력하여 특정 해당 AI 응용 프로그램 또는 구성 요소[25]에 대해 이 섹션에 포함된 요구사항 중 하나 이상을 면제할 수 있으며, 이는 시스템별 위험 평가에 기반하여 요구사항을 충족하면 안전이나 권리에 대한 전반적인 위험이 증가하거나 중요한 기관 운영에 용납할 수 없는 장애를 초래할 수 있다는 서면 결정으로 이루어져야 합니다. 이러한 면제는 AI 사용 기간 동안 적용되지만, AI가 사용되는 조건이나 상황에 중대한 변화가 있는 경우 CAIO가 재평가해야 합니다. 기관 CAIO는 이전에 발급한 면제를 언제든지 취소할 수도 있습니다. 기관은 그러한 면제를 승인한 날로부터 30일 이내에 그 범위, 근거 및 입증 증빙을 자세히 설명하여 OMB에 보고해야 합니다.

iv. 안전에 영향을 미치는 **AI** 또는 권리에 영향을 미치는 **AI**에 대한 최소 관행**.** 2024년 8월 1일부터 기관은 안전에 영향을 미치거나 권리에 영향을 미치는 신규 또는 기존 AI를 사용하기 전에 다음 관행을 따라야 합니다.

A. AI 영향 평가를 완수합니다**.** 영향 평가에서는 다음 사항을 문서화해야 합니다.

1. [AI의 의도된 목적과 기대 이익] 특정 지표 또는 정성적 분석으로 뒷받침되어야 합니다. 지표는 기관의 임무 수행에 긍정적인 결과를 정량화할 수 있는 척도(예: 비용 절감, 고객 대기 시간 단축, 인명 위험 감소 등)여야 하며, AI를 배치한 후 측정하여 AI 사용의 가치를 확인하거나 반증할 수 있어야 합니다[26]. 정량화가 불가능한 경우, 정성적 분석을 통해 고객 경험 개선이나 인간 상호 작용 개선 등 기대되는 긍정적인 결과를 입증하고 AI가 관련 업무를 수행하는 데 적합하다는 사실을 입증해야 합니다.

2. [AI 사용의 잠재적 위험] 더불어 이들 최소 관행 외에 기관이 위험을 감소시키는 데 도움이 되는 추가적인 완화 조치가 있다면 그것이 무엇인지 문서화해야 합니다. 기관은 시스템 사용으로 인해 가장 큰 영향을 받게 될 이해관계자[27]를 문서화하고, AI 및 여타 광범위한 시스템의 오류 가능성 모드에

[25] CAIOs may also grant waivers applicable to groups of closely related AI applications or components, provided that: (1) those systems have undergone a risk assessment that adequately considers the risks from each individual system; and (2) the systems are substantially identical in their risk profiles.

[26] For supervised and semi-supervised AI, agencies should use a target variable which can be reliably measured and adequately represents the desired real-world outcomes.

[27] Stakeholders will vary by use case. For example, if an agency is using AI to control a water treatment process, stakeholders may include (1) local residents; (2) state, local, tribal, and territorial government representatives; and (3) environmental experts.

대하여 평가하되 개별적으로나 인간 사용자 및 시스템 자체의 범위를 벗어난 기타 변수로 인해 발생할 수 있는 결과로 평가해야 합니다. 기관은 특히 소외된 지역사회에 미치는 잠재적 위험에 주의를 기울여야 합니다. AI 기능의 기대 이익은 그 잠재적 위험에 대응하여 고려되어야 하며, 이점이 위험보다 유의미하게 크지 않은 경우 기관은 AI를 사용해서는 안 됩니다.

3. [관련 데이터의 품질 및 적절성] 기관은 AI의 설계, 개발, 학습, 테스트 및 운영에 사용되는 데이터의 품질과 AI의 의도된 목적에 맞는 데이터의 적합성을 평가해야 합니다. 기관이 합리적인 노력을 기울였음에도 이와 같은 데이터에 접근할 수 없는 경우, AI 또는 데이터 제공자에게 이 단락의 보고 요구사항을 준수하는 데 충분한 설명 정보를 구해야 합니다. 최소한 기관은 다음을 문서화해야 합니다.
   a. 의도한 목적에 맞는 데이터의 출처 및 품질[28]
   b. 데이터가 자동화되는 업무와 어떤 관련이 있고 AI의 개발, 테스트 및 운영에 유용할 것이라고 어떻게 합리적으로 기대할 수 있는지
   c. AI가 접할 수 있는 실제 입력 범위를 처리하기에 충분한 범위의 데이터가 포함되어 있는지 여부
   d. 데이터가 충분히 신뢰할 수 있는 출처에서 유래한 것인지 여부
   e. 데이터 입력, 기계 처리 또는 기타 소스에서 발생하는 오류를 어떻게 적절하게 조치하고 제한하는지(학습 데이터 또는 모델 입력에서 AI가 생성한 데이터에 의존하여 발생하는 오류 포함)

B. 실제 상황에서 **AI** 성능을 테스트합니다**.** 기관은 적절한 테스트를 수행하여 AI 및 AI에 의존하는 구성 요소가 의도된 실제 상황에서 작동하는지 확인하여야 합니다. 이러한 테스트는 가능한 경우 분야별 모범 사례를 따라야 하며, 사용된 특정 기술은 물론 시스템 결과에 영향을 미치는 서비스를 사용하는 운영자, 검토자, 직원 및 고객의 피드백을 모두 고려해야 합니다. 테스트 조건은 AI가 배치될 조건과 최대한 유사해야 합니다. 테스트 결과를 통해 기관은 AI가 기대 이익을 달성함과 동시에 AI와 관련된 위험을 충분히 완화할 수 있음을 가능한 범위 내에서 입증해야 합니다. 그렇지 못한 경우 기관은 AI를 사용해서는 안 됩니다. 또한 기관은 강력한 모니터링, 평가 및 안전 조치를 갖춘 파일럿 및 제한적인 배포를 활용하여 광범위한 배포 전에 최종 테스트 단계를 수행하는 것이 바람직합니다.

C. AI에 대하여 독립적으로 평가합니다**.** 기관은 CAIO, 기관 AI 감독 위원회 또는 기존의 테스트 및 평가 책임이 있는 여타의 적절한 기관 사무실을 통해 관련 AI 문서를 검토하여, 시스템이 적절하고 의도한 대로 작동하는지, 기대 이익이 잠재적 위험보다 더 큰지를 확인해야 합니다. 최소한 이 문서에는 섹션 5(c)(iv)에 언급된 실제 상황에서 AI 성능을 테스트한 결과와 완료된 영향 평가가 포함되어야 합니다. 기관은 이 독립적인 평가를 해당 배포, 감독 절차 또는 운영 허가 절차에 통합해야 합니다. 독립적인 검토 기관은 시스템 개발에 직접 관여해서는 안 됩니다.

[28] Consistent with OMB Memorandum M-19-15, Improving Implementation of the Information Quality Act, https://www.whitehouse.gov/wp-content/uploads/2019/04/M-19-15.pdf, and the National Science and Technology Council's report Protecting the Integrity of Government Science, https://www.whitehouse.gov/wpcontent/uploads/2022/01/01-22-Protecting_the_Integrity_of_Government_Science.pdf.

2024년 8월 1일부터 지속적으로, 안전에 영향을 미치거나 권리에 영향을 미치는 신규 또는 기존의 해당 AI를 사용하는 동안, 기관은 AI가 다음 관행을 따르도록 보장해야 합니다.

D. 지속적으로 모니터링을 수행하고 정기 인적 검토의 기준을 수립합니다**.** 배치 전 테스트 외에도 기관은 AI 기능 저하를 모니터링하고 AI가 권리나 안전에 미치는 영향의 변화를 감지할 수 있는 지속적인 절차를 마련해야 합니다. 이 모니터링 절차의 일부에는 이 섹션 최소 관행의 기존 구현이 신규 위험을 적절하게 완화하는지 여부를 확인하기 위하여 정기적인 인적 검토를 포함해야 합니다. 실제 상황에서 AI 성능에 대한 새로운 테스트를 포함하여 이러한 인적 검토는 적어도 매년[29], 또는 AI나 AI가 사용되는 조건이나 상황이 크게 변경된 경우 수행되어야 합니다. 검토에는 시스템 개발이나 운영에 직접 관여하지 않는 적절한 기관 내부 기구의 감독 및 검토가 포함되어야 합니다. 또한 기관은 가능한 경우 새롭거나 업데이트된 AI 기능의 사용을 점진적으로 확대하여 부정적인 성능이나 결과를 모니터링할 수 있는 충분한 시간을 확보해야 합니다. 또한 기관은 특히 권리나 안전에 부정적인 영향을 미칠 수 있는 AI 특화 공격[30]으로부터 AI를 모니터링하고 방어해야 합니다.

E. 권리와 안전에 미치는 신규 위험을 완화합니다**.** 지속적인 모니터링, 정기 검토 또는 기타 메커니즘을 통해 권리 또는 안전에 미치는 새롭거나 상당히 변경된 위험을 파악한 경우, 기관은 해당 위험을 완화하기 위한 조치를 취해야 하며, 여기에는 적절한 경우 AI를 업데이트하여 위험을 줄이거나 인적 감독 강화 등 비기술적 완화 조치를 시행하는 것이 포함됩니다. 상당한 변경으로 인하여 학습이나 문서가 부정확해지는 등 기존에 이 섹션에서 설명한 여타 최소 관행을 이행한 효과가 떨어지는 경우, 기관은 해당 관행을 적절하게 업데이트하거나 반복해야 합니다. 권리 또는 안전에 미치는 AI의 위험이 용납할 수 있는 수준을 넘어서고 완화 조치가 실행 불가능한 경우, 기관은 가능한 한 빨리 해당 영향을 받는 AI의 사용을 중단해야 합니다[31].

F. 인력에 대한 적절한 교육훈련 및 평가를 보장합니다**.** 기관은 AI 운영자에 대하여 충분한 교육, 평가 및 감독을 보장하여 이들이 AI의 결과를 해석하고 이에 따라 조치를 취하며, 인간-기계 팀의 구성 문제(예: 자동화 편향)에 대처할 수 있고, 시스템의 인간 기반 구성 요소가 AI 활용으로 인한 위험을 효과적으로 관리할 수 있도록 보장해야 합니다. 교육훈련은 해당 기관의 결정에 따라 주기적으로 실시되어야 하며, 운영 중인 AI 사용 사례, 제품 및 서비스에 맞게 구체적이어야 합니다.

G. 권리나 안전에 고위험을 초래하는 결정의 일부로서 적절한 인적 검토를 제공합니다**.** 기관은 권리나 안전에 고위험을 초래하는 결정 과정에서 역할을 하는 AI

[29] For customer-facing services, agencies should consider customer feedback.

[30] For example, the AI-specific exploits outlined in the MITRE ATLAS framework. See https://atlas.mitre.org/.

[31] Agencies are responsible for determining how to safely decommission AI that was already in use at the time of this memorandum's release without significant disruptions to essential government functions.

기능을 파악하고, 적절한 인적 검토와 책임성 없이는 AI 기능이 그러한 상황에 직접 개입하지 못하도록 조치해야 합니다.

H. AI 사용 사례 인벤토리를 이용하여 공개적으로 공지하고 평이한 언어로 작성된 문서를 제공합니다**.** 기관은 개인정보 보호, 민감한 법 집행, 국가 안보 및 기타 보호 정보에 관한 지침을 포함하여 관련 법률 및 정부 차원의 지침에 부합하는 범위 내에서, AI의 사용 사례 인벤토리가 시스템 기능에 대하여 적절하고 상세하며 일반적으로 이용될 수 있는 문서로서 제공되어, 사용자와 일반 대중에게 AI에 대한 공개 공지가 될 수 있도록 보장해야 합니다. 가능한 경우 기관은 사람들이 AI와 상호작용하거나 영향을 받을 수 있는 상황에 이 문서를 포함하거나 링크를 제공해야 합니다. 기관의 사용 사례가 본 지침에 설명된 공개 인벤토리 요구사항에서 제외되는 경우에도 해당 기관은 관련 정보를 OMB에 보고해야 할 수 있으며, 관련 법률에 따라 적절하고 일관성 있게 AI 사용에 있어 적합한 투명성을 보장해야 합니다.

v. 권리에 영향을 미치는 **AI**에 대한 추가적인 최소 관행

2024년 8월 1일부터 기관은 상기의 안전에 영향을 미치거나 또는 권리에 영향을 미치는 AI에 대하여 상기 최소 관행 중 하나를 따라야 합니다. 추가적으로 기관은 권리에 영향을 미치는 신규 또는 기존 AI를 사용하기 전에 다음과 같은 최소 관행도 따라야 합니다.

A. AI가 형평성**,** 존엄성**,** 공정성을 증진할 수 있도록 조치합니다**.** 여기에는 최소한 다음 조치가 포함되어야 합니다.

1. [알고리즘 차별 또는 편향에 기여하는 요인을 사전에 파악하고 제거] 기관은 권리에 영향을 미치는 기관 AI가 연방 차별금지법이 보호하는 계층에 대한 알고리즘 차별이나 편향을 초래할 수 있는 방식으로 해당 계층에 대한 정보에 실질적으로 의존하는지 평가해야 합니다. 기관은 또한 대리변수가 권리에 영향을 미치는 기관 AI에 과도한 영향력을 만드는지 여부를 평가해야 합니다. 두 경우 모두, AI가 그러한 정보에 의존하여 보호 대상 계층에 대한 불법적인 차별이나 유해한 편견을 초래하는 경우, 기관은 AI를 의사결정에 사용하기 전에 해당 정보의 사용을 중단해야 합니다.

2. [서로 다른 영향의 평가 및 완화] 기관은 기관 AI의 실제 배치를 비롯하여 인구통계학적 집단 전반에서 AI 성능에 상당한 격차가 발생하는지 여부를 확인하기 위하여 AI를 테스트해야 합니다. 관련 법률에 따라, 차별을 초래할 가능성이 있거나, 중대한 피해를 야기하거나, 또는 형평성, 존엄성 또는 공정성을 저해할 가능성이 있는 격차를 적절히 해결해야 합니다. 격차를 적절히 완화할 수 없는 경우 기관은 AI 도구를 사용하거나 통합해서는 안 됩니다.

3. [대표 데이터의 사용] 기관은 기관 AI를 개발, 운영, 평가하는 데 사용되는 데이터가 AI의 영향을 받게 될 집단을 적절하게 대표하였는지 확인하고, 데이터의 역사적, 사회적 상황으로 인한 부적절한 편향성을 검토했는지 확인해야 합니다.

B. 영향을 받는 집단의 피드백을 수렴하고 반영합니다**.** 관련 법률 및 정부 차원 지침에 따라 실행 가능한 범위 내에서, 기관은 AI를 설계, 개발 및 사용할 때 소외 계층을 포함하는 영향을 받는 집단의 의견을 수렴하고, 이러한 피드백을 AI에 관한 기관의 의사 결정 과정에 제보해야 합니다. 부정적인 피드백이 있을 경우, 기관은 AI를 배치하지 않거나 AI 사용을 중단할 것을 고려해야 합니다. 기관이 고객[32], 연방 직원 모임 및 노동조합 대표자 등 영향을 받는 집단에 지속적으로 피드백을 요청할 것을 강력히 권장하며, 특히 AI나 AI가 사용되는 조건이나 상황이 크게 변경된 후에 더욱 그렇습니다. 이러한 의견수렴을 실시하기 위하여 기관은 AI의 영향을 받는 집단에게 의견을 받기 위한 적절한 조치를 취해야 하며, 여기에는 다음이 포함될 수 있습니다[33].

1. 시스템과 상호 작용하는 사용자의 관찰 등 직접적인 사용자 테스트

2. 일반 대중의 의견을 일반적으로 수렴하는 절차. 연방 관보에 대한 정보 요청 또는 응답을 개방적으로 기입할 수 있는 "귀하의 경험을 들려주십시오" 지면 등

3. 서비스 후 고객 피드백 수집[34]

4. 공청회 또는 회의. 청중 토론 등

5. 영향을 받는 집단에게 유의미하고 공평하며 접근 가능하고 효과적인 방식으로 공개적인 참여, 의견 또는 피드백을 구하는 기타 투명한 절차

2024년 8월 1일부터 지속적으로, 권리에 영향을 미치는 신규 또는 기존의 해당 AI를 사용하는 동안, 기관은 AI가 다음 관행을 따르도록 보장해야 합니다.

C. AI 기반 차별에 대하여 지속적으로 모니터링하고 완화를 실시합니다**.** 섹션 5(c)(iv)(D)에 언급된 지속적인 모니터링 요구사항의 일환으로, 기관은 권리에 영향을 미치는 AI를 모니터링하여, 예상치 못한 상황, 배치 후 시스템 변경, 사용 상황 또는 관련 데이터 변경으로부터 초래될 수 있는, 보호 대상 계층에 대한 AI 기반 차별을 평가하고 완화하여야 합니다. 충분한 완화가 불가능할 경우, 기관은 영향을 받는 AI 기능의 사용을 안전하게 중단해야 합니다.

D. 부정적인 영향을 받은 개인에 대하여 통지합니다**.** 관련 법률 및 정부 차원 지침에 따라 실행 가능한 경우, AI가 급여 지급 거부 등 개인들과 구체적으로 관련된 의사

[32] Customers can include individuals, businesses, or organizations that interact with an agency.

[33] Agencies are not required to conduct consultations in a format that would require OMB clearance under the Paperwork Reduction Act (44 U.S.C. § 3507), provided the steps the agency takes are adequate to solicit input from the groups affected by the AI.

[34] Information on post-transaction customer feedback surveys can be found in OMB Circular A-11, Section 280 – Managing Customer Experience and Improving Service Delivery, https://www.whitehouse.gov/wpcontent/uploads/2018/06/s280.pdf.

결정 결과에 중대한 영향을 미치게 될 때 기관은 개인들에게 이를 알려야 합니다[35]. 이러한 통지는 적시에 이루어져야 하며 해당되는 경우 2010년 평이한 문서작성법(Plain Writing Act)[36]에 부합하는 방식으로 작성되어야 합니다. 기관은 AI 사용 상황에 따라 통지 시기는 물론 대체 형식 및 채널을 통해 다중언어로 통지하는 데 적절한 시기를 고려해야 합니다. 또한 통지에는 해당 기관에 연락할 수 있는 명확하고 이용 가능한 수단이 포함되어야 하며, 적절한 경우 관련 문제에 대하여 시기적절한 해결을 요청할 수 있어야 합니다. 또한 기관이 그러한 결정과 조치에 대해 설명을 제공할 것을 강력히 권장합니다[37].

E. 인적 검토와 구제 절차를 관리합니다**.** 영향을 받은 개인이 자신에 대한 AI의 부정적인 영향에 대하여 항고하거나 이의를 제기하려는 경우, 기관은 대체 시스템 및 단계적 확대 시스템을 통해 AI 사용에 대한 인적 검토와 잠재적 구제수단을 적시에 제공해야 합니다. 적절한 구제수단을 개발할 때 기관은 행정 부담 산정에 대한 OMB 지침을 따라야 하며 구제 절차가 영향을 받은 개인에게 불필요한 부담을 주어서는 안 됩니다[38]. 법률 또는 정부 차원의 지침에 따라 AI 사용 또는 개인 항고 기회에 대한 공개가 배척되는 경우, 기관은 권리에 영향을 미치는 AI를 인간이 감독할 수 있는 적절한 메커니즘을 마련해야 합니다.

F. 가능한 경우 거부**(opt-out)** 옵션을 관리합니다**.** 기관은 AI 기능을 편리하게 거부할 수 있는 메커니즘을 명시적으로 제공하고 관리하여야 하며, 실행 가능하고 관련 법률 및 정부 차원의 지침에 부합하는 경우 인간의 대안을 선호해야 합니다. 영향을 받는 사람들이 대안에 대하여 합리적인 기대를 가지고 있거나, 대안이 부족하여 접근성을 중대하게 제한하거나 부당하게 유해한 영향을 미치는 경우, 거부 메커니즘이 반드시 존재해야 합니다.

d. 연방 정부 인공지능 조달의 위험 관리

이 섹션에서는 책임 있는 연방 정부 AI 조달에 대한 권장 사항을 제시합니다. 이들 권장 사항 외에도 미국 AI 발전법 섹션 7224(d) 및 AI 행정 명령 섹션 10. 1(d)(ii)에 따라, OMB는 AI 계약이 이 각서의 지침에 부합하도록 보장할 수 있는 일차 수단도 개발할 예정입니다.

---

[35] In some instances, such as an active law enforcement investigation, providing immediate notice may be inappropriate or impractical, and disclosure may be more appropriate at a later stage (i.e., prior to a defendant's trial).

[36] Pub. L. No. 111-274 (codified at 5 U.S.C. § 301 note), https://www.congress.gov/111/plaws/publ274/PLAW111publ274.pdf.

[37] Explanations might include, for example, how and why the AI-driven decision or action was taken. While exact explanations of AI decisions are often not technically feasible, agencies should characterize the general nature of such AI decisions through context such as the data that the decision relied upon, the design of the AI, and the broader decision-making context in which the system operates. Such explanations should be technologically valid, meaningful, useful, and as simply stated as possible, and higher-risk decisions should be accompanied by more comprehensive explanations.

[38] See OMB M-22-10 and supporting document "Strategies for Reducing Administrative Burden in Public Benefit and Service Programs."

i. 국가 가치 및 법률의 준수**.** 기관은 조달된 AI가 국가의 가치에 적합한 존중을 나타내고, 헌법에 합치하며, 개인정보 보호, 기밀 유지, 저작권, 인권 및 시민권, 시민 자유를 다루는 여타의 모든 관련 법률, 규정 및 정책을 준수하도록 보장해야 합니다.

ii. 투명성 및 성능의 개선**.** 기관은 조달된 AI의 투명성과 적절한 성능을 보장하기 위하여 다음과 같은 조치를 취해야 합니다.

A. 모델, 데이터, 시스템 카드 사용 등을 통해 조달된 AI에 대한 적절한 문서를 확보합니다.

B. 기관이 기능을 배치할 것으로 예상되는 특정 환경을 포함하여, 연방 계약업체가 제기한 AI 성능에 대한 주장을 정기적으로 평가합니다.

C. 조달된 AI의 지속적인 개선을 장려하는 계약 조항을 고려합니다.

iii. AI 조달의 경쟁 촉진**.** 기관은 연방 AI 조달 관행이 계약업체 간의 경쟁 기회를 촉진하고 기존 업체에 부적절하게 고착되지 않도록 적절한 조치를 취해야 합니다. 이러한 조치에는 상호 운용성을 촉진하고 공급업체가 경쟁업체의 제품을 희생하면서 자사 제품을 부적절하게 선호하지 않도록 보장하는 조치가 포함될 수 있습니다.

iv. AI를 위한 데이터 가치의 극대화**.** AI 제품 및 서비스 계약에 있어 기관은 관련 데이터는 물론 해당 데이터에 대한 수정(예: 정제 및 라벨링)을 AI 성숙도를 위한 중요한 자산으로 취급해야 합니다. 기관은 계약이 데이터 및 해당 데이터의 개선 사항에 대한 충분한 권리를 정부가 보유하도록 조치를 취하여, 공급업체 종속(lock-in)을 방지하고 정부의 지속적인 AI 설계, 개발, 테스트 및 운영을 촉진해야 합니다. 또한 기관은 연방 정부 AI 제품 및 서비스의 개발 및 운영 과정에서 공급업체가 사용하는 연방 정보를 보호하는 계약 조항을 고려하여, 해당 데이터가 무단으로 공개되거나 사용되지 않고 이후 기관의 명시적인 허가 없이는 공급업체가 제공하는 상업용 AI 제품의 기능을 학습시키거나 개선하는 데 사용할 수 없도록 보호하는 조치를 취해야 합니다.

v. 책임 있는 생성형 **AI** 조달**.** 기관은 생성형 AI, 특히 이중 용도 파운데이션 모델에 대한 계약에서 다음을 포함하는 맞춤형 위험 관리 요구사항을 포함하는 것이 바람직합니다.

A. 차별적이거나, 호도하거나, 선동적이거나, 안전하지 않거나 기만적인 결과물 등 생성형 AI로 인한 위험에 대응하는 외부 레드티밍을 비롯하여, 적절하게 테스트하고 안전 조치를 취할 것을 요구합니다.

B. 적절하고 기술적으로 실행 가능한 경우, 생성형 AI 모델이 AI가 생성하거나 수정한 콘텐츠에 대하여 출처를 안정적으로 표시하거나 입증할 수 있는 기능을 갖출 것을 요구합니다.

C. 기관은 이러한 요구사항을 부과할 때 AI 행정 명령의 섹션 4. 1(a) 및 10. 1(d)에 따라 정의된 관련 NIST 표준을 적절하게 고려하는 것이 바람직합니다.

6. 정의

(이하 생략)