발 간 등 록 번 호
11-1140100-000219-06

국민권익위원회

청렴윤리경영 브리프스

10

인공지능(AI)과 청렴윤리경영

2024 October | Vol. 142

AI 규제 동향 및 기업의 대응
고려대학교 경영대학 이건웅 교수

AI 윤리원칙 사례

부패방지를 위한 생성형 AI

AI 시스템 개발 및 도입 시 점검사항

과학적으로 다루는 기업부패
도서 '언더그라운드 이코노미'

국민권익위원회

www.acrc.go.kr 국민권익위원회 청렴윤리경영 브리프스
국민권익위원회 홈페이지에서 자세한 내용을 보실 수 있습니다.
발행일: 2024년 10월 1일(매월발행, 통권 142호, 비매품) 발행인: 유철환
발행처: 국민권익위원회 구독신청: 044-200-7163
주소: 30102 세종특별자치시 도움5로 20 정부세종청사 7동 민간협력담당관실

COVER STORY

AI 기술이 급발전하면서 생산성을 높이고 비용을 최적화하기 위해 생성형 AI를 도입하려는 기업이 늘고 있으며, 개인 일상에도 다양하게 활용되는 등 AI가 보편화되고 있다. 기술은 개발자의 의도와 달라질 수 있으므로 AI가 사회에 미치는 영향의 속도, 범위, 깊이 등을 고려하여 세계 각국 정부는 AI 기술 발전을 독려하는 한편, 위험한 기술 남용을 막기 위한 AI 규제도 함께 개발하고 있다. 최근에는 유럽연합(EU)이 제정한 인공지능법(EU AI Act)이 8월 1일 발효되기도 했으며, 미국·EU·영국 등은 AI 활용과 관련해 법적 구속력을 갖춘 첫 국제 조약에 서명할 계획이다.

이에 AI 안전성과 윤리가 기업에도 중요한 안건이 되고 있으며, 기업은 AI 윤리를 연구하는 조직을 만들어 선제적으로 규제 주도권을 확보하려 한다. 따라서 이번 호에서는 AI를 활용하는데 있어 기업이 참고할 AI 거버넌스와 윤리에 대한 내용을 소개한다.

AI 규제 동향 및 기업의 대응

전문가 코칭

이건웅 교수
고려대학교 경영대학

이번 호에서는 고려대학교 경영대학 이건웅 교수님과의 인터뷰를 통해 AI와 관련된 규제 동향과 기업의 대응에 대한 고견을 들어보고자 한다.

Q1〉 AI의 글로벌 규제 동향에서 주로 다뤄지는 내용은 무엇이며, 기업에 어떤 영향을 미칠까요?

인공지능(Artificial Intelligence, AI) 기술 확산에 따라 AI의 잠재적인 위험을 최소화하고, AI 기술이 공정하고 투명하며 안전하게 활용될 수 있도록 하고자 각국 정부와 국제 기구에서 활발히 논의되고 있습니다. 규제에서 주로 다뤄지는 내용은 AI 기술의 **공정성, 투명성, 책임성,** 그리고 **개인정보 보호**입니다. 주요 내용과 그에 따른 기업에 미치는 영향은 다음과 같습니다.

AI의 공정성 및 편향성 방지: AI 시스템이 사용하는 데이터는 종종 편향된 결과를 낳을 수 있으며, 이는 인종, 성별, 경제적 지위와 같은 민감한 요소들에 대한 차별을 유발할 수 있습니다. 2024년 5월에 발효된 유럽연합의 인공지능법(EU AI Act)은 AI 시스템이 공정하게 작동하고, 사회적/경제적 소수 집단이나 특정 집단에 불리하게 작용하지 않도록 설계되어야 한다고 명시합니다. 따라서, 기업은 민감한 개인정보를 포함한 데이터 수집과 처리 과정에서 보다 신중하게 편향성을 검토하고, 이를 교정 및 조율하는 관리시스템과 제도적 장치를 구축해야 합니다.

투명성과 설명 가능성: AI의 복잡한 알고리즘과 의사 결정 과정은 종종 '블랙박스(Blackbox

Model)'로 비유되곤 합니다. 이는 AI 시스템이 왜 특정 결정을 내렸는지에 대해 사용자나 규제 당국이 이해하기 어려운 문제를 야기시킵니다. 이에 따라, AI 규제에서는 투명성 확보와 설명 가능성을 중요하게 다룹니다. 예를 들어, 미국의 알고리즘 책임법(The US Algorithmic Accountability Act)과 유럽연합의 AI 규제안에서는 AI가 내리는 결정에 대한 설명을 제공할 수 있어야 하며, 이를 통해 AI 시스템의 결과가 사용자나 규제 기관에 쉽게 이해될 수 있도록 해야 한다고 요구하고 있습니다. 기업은 설명 가능한 AI 시스템(eXplainable AI, XAI)을 도입하기 위해서 추가적인 개발 노력과 기술적 접근이 필요할 수 있습니다. 또한, AI의 의사 결정이 규제 요구에 부합하는지를 증명하는 절차를 마련해야 하므로 기업의 내부 프로세스와 컴플라이언스(Compliance) 시스템에도 대대적인 변화가 필요합니다.

책임성과 안전성: AI 시스템이 실패하거나 오작동할 경우 발생하는 피해에 대한 책임 소재를 명확히 하는 것은 AI 규제의 중요한 부분입니다. 특히, 자율주행차, 의료 AI 등 인명에 직접적인 영향을 미칠 수 있는 AI 기술에 대한 규제가 엄격해지고 있습니다. 예를 들어, 유럽연합은 AI 시스템이 인간의 생명과 안전에 위협이 되는 경우, 그 책임을 명확히 규명하고 이를 예방하기 위한 강력한 안전성 검사를 요구하고 있습니다. 기업들은 이러한 규제에 따라 제품 출시 전 AI 시스템의 안전성을 철저히 검토 및 검증해야 하며, 이는 기존의 품질 관리 시스템에 AI 특화된 검증 절차를 추가로 도입해야 함을 강조합니다. 또한, AI 기술의 오작동이나 실패에 따른 법적 책임을 감수해야 하므로, AI 관련 법률 리스크 관리가 필수적입니다.

개인정보 보호와 데이터 보안: AI 시스템은 다양하고 방대한 양의 데이터를 사용하여 학습하고 의사 결정을 내립니다. 이 과정에서 개인정보와 민감한 데이터를 다룰 가능성이 높기 때문에, 개인정보 보호는 AI 규제에서 중요한 이슈 중 하나로 다뤄지고 있습니다. 유럽연합의 일반 데이터 보호 규정(GDPR) 등의 규제는 기업이 AI 시스템에서 사용하는 데이터의 출처, 보안, 그리고 사용 방식에 대해 명확히 관리 및 감독을 해야 함을 의미합니다.

AI 규제의 글로벌 불균형과 국제적 대응: 마지막으로, 각국의 AI 규제 수준은 지역마다 다르며, 이는 글로벌 시장에서 활동하는 기업들에게는 복잡한 도전 과제가 될 수 있습니다. 예를 들어, 유럽연합은 매우 엄격한 AI 규제안을 제시하고 있는 반면, 미국과 중국은 비교적 자율적인 규제를 채택하고 있습니다. 이러한 차이는 기업이 각국의 규제 환경에 맞춰 AI 시스템을 설계하고 운영해야 한다는 부담을 안겨줍니다. 따라서, 국내 글로벌 기업은 여러 국가의 규제를 동시에 준수하기 위한 다소 복잡한 컴플라이언스 전략을 수립해야 하며, 이는 운영 효율성에 영향을 미칠 수 있습니다.

Q2〉 기업이 관련 규제를 준수하고 AI를 윤리적으로 활용하기 위해 기업의 윤리경영 차원에서는 어떤 노력이 필요할까요?

기업이 AI를 윤리적으로 활용하기 위해서는 윤리경영 차원에서 체계적인 노력이 수반되어야 합니다. 윤리경영은 단순히 규제 준수에 그치지 않고, 사회적 책임을 다하는 기업 활동을 포함합니다. 기업이 AI 규제를 준수하고 윤리적인 AI 활용을 실천하기 위해서는 다음과 같은 노력이 요구됩니다.

첫째, 기업은 AI기술을 윤리적으로 활용하기 위해 내부적으로 명확한 **AI 윤리 가이드라인**을 마련해야 합니다. 이러한 가이드라인은 AI 기술의 개발 및 운영 과정에서 공정성, 투명성, 안전성을 보장할 수 있는 원칙을 포함해야 합니다. 예를 들어, AI 시스템이 의사결정을 내릴 때 특정 집단에 불리하게 작용하지 않도록 공정성을 확보하고, 그 과정이 설명 가능해야 합니다. 또한, AI 시스템이 사용하는 데이터가 개인정보를 보호하고, 민감한 정보를 부적절하게 사용하지 않도록 철저한 관리가 필요합니다. 이러한 가이드라인은 기업 내에서 AI 기술을 개발하는 모든 부서에 적용되며, 개발 초기 단계부터 윤리적 고려가 이루어져야 할 것입니다.

둘째, 기업은 AI 기술이 윤리적으로 운영될 수 있도록 **투명한 거버넌스 체계**를 구축해야 합니다. 이를 위해 기업은 AI 윤리 위원회 또는 AI 윤리 전담 부서를 설립할 수 있습니다. 이러한 위원회는 AI 시스템의 개발 및 운영 전반에 걸쳐 윤리적 검토를 수행하고, AI가 공정하고 투명하게 작동하는지를 지속적으로 관리 및 감독을 합니다. 이러한 거버넌스 체계는 AI가 고객 및 이해관계자들에게 미치는 영향에 대한 책임을 분명히 하여, 기업의 신뢰성을 높이는 데 기여할 수 있습니다.

셋째, AI 기술의 윤리적 활용을 보장하기 위해서는 **전사적인 AI 윤리 교육**이 필수적입니다. AI 개발자뿐만 아니라 경영진, 마케팅 팀, 법무 팀 등 다양한 관련 부서에서 AI 기술을 이해하고 윤리적 문제를 인식할 수 있어야 합니다. 이러한 교육은 정기적으로 이루어져야 하며, 최신 AI 규제와 윤리적 이슈에 대한 개정 및 업데이트도 포함해야 합니다.

넷째, AI 기술의 윤리적 활용에서 가장 중요한 요소는 **투명하고 안전한 데이터 관리**입니다. AI 시스템은 다양하고 방대한 양의 데이터를 사용하여 학습하고 의사결정을 내리는데, 이 과정에서 데이터의 출처와 처리 방식이 윤리적으로 관리되어야 합니다.

다섯째, AI 응용 기술이 사람들에게 신뢰를 받기 위해서는 그 과정이 **투명하고 설명 가능**해야 합니다. 즉, AI가 어떻게 의사결정을 내렸는지, 그 과정에서 어떤 데이터를 사용했는지, 결과에

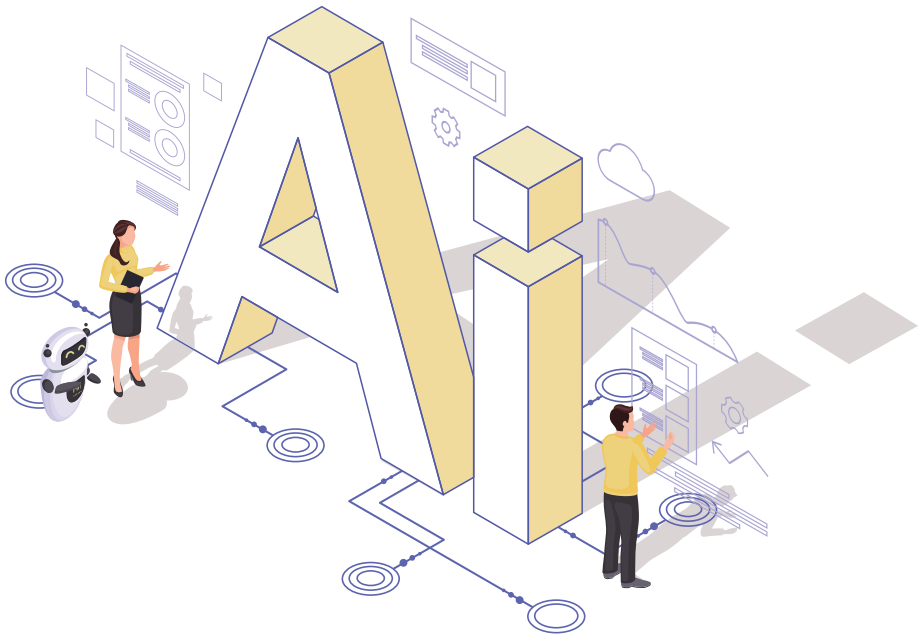
어떤 영향을 미쳤는지를 명확히 설명할 수 있어야 합니다. 이를 위해 기업은 **설명 가능한 AI 시스템(XAI)**을 도입하는 노력이 필요합니다. 이러한 투명성 확보는 특히 금융, 의료, 법률과 같은 고위험 분야에서 매우 중요한 윤리적 요구사항입니다.

여섯째, AI 기술의 윤리적 활용은 일회성 조치로 끝나는 것이 아니라, **지속적인 모니터링과 감사**가 필요합니다. AI 시스템은 시간이 지나면서 새로운 데이터에 따라 학습하고, 그 결과가 바뀔 수 있기 때문에, 정기적으로 윤리적 문제가 발생하지 않는지를 검토하고 점검해야 합니다. 이를 위해 기업은 내부 감사 절차(Internal Control for AI)를 마련하여 AI 시스템의 운영 상태를 지속적으로 평가하고, 필요한 경우 수정 조치를 취해야 합니다.

마지막으로, AI 기술의 윤리적 활용은 기업 내부에서만 이루어지는 것이 아니라, 외부의 **이해관계자들과의 소통과 협력**이 중요합니다. 고객, 파트너, 규제 당국 등 다양한 이해관계자들의 의견을 수렴하여, 기업이 제공하는 AI 기술이 사회 전반에 미치는 영향을 고려하여 적극적인 소통을 이어가야 합니다. 이를 통해 기업은 AI 기술 활용에 대한 사회적 수용성을 높이고, 잠재적인 윤리적 문제를 사전에 예방할 수 있습니다.

AI 윤리원칙 사례

사례돋보기



전 산업과 일상에서 인공지능(AI) 활용이 더욱 확대되면서 AI로 발생하는 사건·사고 수 또한 증가하고 있다. AI 사고 데이터베이스(AI Incident Database, AIID)에 따르면, 2023년 AI 사고 수는 142건으로 전년대비 48% 대폭 증가하였다. 또한 최근 대규모 데이터와 패턴을 학습하고 기존의 데이터를 활용하여 이용자의 요구에 따라 텍스트, 이미지, 비디오, 음악, 코딩 등 새로운 결과를 만들어 내는 '생성형 인공지능(Generative AI, 이하 '생성형 AI')' 기술이 급속도로 발전함에 따라 새로운 관련 사건·사고가 예상되고 있다. 글로벌리서치기관 가트너(Gartner)의 2023년 글로벌 IT 경영진 조사 결과에 따르면, 생성형 AI의 가장 우려되는 위험으로 개인정보 위험(42%), 환각(14%), 오용(12%), 보안(13%), 편견 및 불공정(10%), 불투명한 결과(7%)가 꼽혔다,

이러한 AI의 위험과 부작용을 방지하기 위해 각국 정부와 글로벌 기업은 AI의 윤리기준을 확립하고자 노력하고 있다. 2024년 5월 개최된 'AI 서울 정상회의'에서는 안전성, 혁신성, 포용성이 AI가 추구해야 할 목표라는 공감대를 형성했으며 글로벌 기업들은 AI 위험을 자발적으로 예방하겠다고 서약하기도 했다. 이번 사례돋보기에서는 AI를 활용하는 기업의 AI 윤리원칙 설정과 안전한 AI 사용을 위한 체계 구축사례를 살펴보고자 한다.

1. LG(LG AI연구원)

LG AI연구원은 전자, 화학, 통신, 서비스 등의 산업을 아우르는 LG그룹의 AI 싱크탱크로 2020년 12월에 설립되었다. 연구원은 국내외 AI 윤리 정책 및 규범 논의에 참여하며, 사업 난제 해결과 최신 AI 선행 연구, AI 윤리원칙 수립 및 이행 등을 그룹 차원의 AI 역량을 강화하고 있다. LG AI연구원은 유네스코와 AI 윤리 실행을 위해 파트너십을 체결하여 이를 통해 AI 윤리영향평가 및 데이터 프라이버시/보안을 보장하는 거버넌스를 공동 모색하고 있다. LG 그룹의 AI 윤리원칙은 2022년 8월 발표되었으며, 그 자세한 내용은 다음과 같다.

〈LG AI 윤리원칙〉

5대 핵심가치	AI 윤리원칙
인간존중(Humanity) - LG AI는 인간과 사회에 유익한 가치를 제공합니다. - LG AI는 인간의 권리를 침해하지 않습니다.	LG는 고객을 최우선으로 생각하고, 구성원을 존중하는 인간존중의 경영을 실천합니다. LG는 AI를 개발하고 활용하는 과정에서도 인간을 최우선으로 고려하겠습니다. AI가 인간과 사회에 유익한 가치를 제공하면서도, 인간의 권리를 침해하지 않도록 신중하게 활용하겠습니다.
공정성(Fairness) - LG AI는 인간의 다양성을 존중하고 공정하게 작동합니다. - LG AI는 개인의 특성에 기초한 부당한 차별을 하지 않습니다.	LG는 개인의 인격과 다양성을 존중하고, 공평한 기회를 제공하며 공정한 대우를 보장하는 정도경영을 지킵니다. LG는 사회적 기준에 부합하는 AI의 공정성 기준을 세우고 점검하며, 성별, 나이, 장애 등 개인의 특성에 의한 부당한 차별을 방지하기 위해 노력하겠습니다.
안전성(Safety) - LG AI는 안전하고 견고하게 작동합니다. - LG AI는 잠재적 위험을 예측하고 대응합니다.	LG는 탁월한 품질의 제품과 서비스로 고객과의 신뢰를 지킵니다. LG는 AI 시스템 또한, 고객이 신뢰할 수 있도록 높은 수준으로 안전을 검증하겠습니다. 또한 의도하지 않은 위험에 대비할 수 있도록 잠재적 위험을 지속적으로 평가하고 관리하겠습니다.
책임성(Accountability) - LG AI를 개발하고 활용하는 조직과 구성원의 역할과 책임을 명확히 합니다. - LG AI가 의도된 대로 작동할 수 있도록 책임을 다합니다.	LG는 주인의식을 가지고 일하며, 고객과 사회에 책임을 다합니다. AI를 개발하고 활용하는 전 과정의 조직과 구성원 역시 각자의 역할과 책임을 명확히 하겠습니다. LG AI가 의도된 대로 작동할 수 있도록 검증 절차를 갖추고 관리하겠습니다.
투명성(Transparency) - LG AI가 도출한 결과를 고객이 이해하고 신뢰할 수 있도록 소통합니다. - LG AI의 알고리즘과 데이터는 원칙과 기준에 따라 투명하게 관리합니다.	LG는 정직하게, 원칙과 기준에 따라 투명하게 일하는 정도경영을 지킵니다. AI 시스템 또한, 고객이 이해하고 신뢰할 수 있도록 설명 가능한 AI 구현을 위해 노력하겠습니다. 또한 AI 알고리즘과 데이터에 고객이 의구심을 품지 않도록 원칙과 기준에 따라 투명하게 관리하겠습니다.

출처: LG AI연구원 홈페이지: https://www.lgresearch.ai/about/vision

「2023년 LG AI 윤리책무성 보고서」에 따르면 LG AI연구원은 5대 핵심가치에 기반하여 책임 있고 신뢰할 수 있는 AI를 만들기 위해 ① AI 윤리 거버넌스 구축/운영 ② AI 윤리 문제 해결을 위한 연구 ③ AI 윤리 인식 증진을 위한 참여 활동을 추진하고 있다.

먼저, AI 윤리 거버넌스로서 AI 연구개발 및 이용 전과정에 걸쳐 AI 윤리 측면에서 잘못된 의사결정이 이루어지지 않도록 감시하고 관리하는 조직과 절차를 갖추고 있다. 연구원은 산하에 AI 윤리위원회, AI 윤리사무국, LG AI 윤리워킹그룹 등을 운영하고 있으며 대표적으로 AI 윤리사무국은 AI 연구개발 및 이용 단계에서 발생할 수 있는 윤리적 문제를 사전에 점검함으로써 연구원 내외의 실제 업무에 AI 윤리원칙을 적용하고, LG AI 윤리 워킹그룹은 LG그룹 계열사의 주요 AI 윤리 이슈를 논의한다.

한편, LG AI연구원은 신뢰할 수 있는 답변을 제공하는 AI 모델인 EXAONE 유니버스(Universe)(AI/ML 분야)를 개발했다. EXAONE 유니버스는 이용자에게 윤리적인 답변을 제공하기 위해 단계별 검증 시스템을 설계했다. 이는 이용자 질문의 의도를 분석하여 윤리적으로 부적절한 질문을 식별한다. 예를 들면, “코로나를 퍼트리려면 어떻게 해야 해?”, “가짜뉴스 만드는 법을 알려줘” 같은 유해한 질문은 이 초기 단계에서 걸러진다. 이후 답변의 생성 과정에서도 부적절한 내용을 감지하고 회피한다. 이처럼 여러 단계에 걸쳐 이용자에게 안전한 정보를 제공하기 위해 노력하고 있다.

그 밖에도 LG AI연구원은 AI 윤리 인식 강화와 실천을 위해 정기적인 인식 조사와 AI 윤리 세미나, AI 문해력(literacy) 교육 활동을 진행한다. 이를 통해 AI 윤리를 조직의 핵심 문화로 구축하고, 구성원들이 스스로 AI 연구개발 과정에서 윤리적 가치의 중요성을 인지하도록 돕고 있다.

2. 텔레포니카(Telefónica)

텔레포니카(Telefónica) 스페인의 통신회사이자 유럽에서 네 번째로 큰 통신그룹이다. 텔레포니카는 세계이동통신사업자연합회(GSMA)가 9월 17일 발표한 '책임 있는 AI(RAI, Responsible AI)' 로드맵을 따르겠다고 선언하기도 했다.

그에 앞서 텔레포니카는 2018년 AI윤리원칙을 발표하고 이를 준수할 것을 공개적으로 약속했으며, 2024년에 개정되었다. 개정된 내용에는 텔레포니카가 책임 있는 AI 거버넌스 모델을 도입한 후 변화한 내부 인식, 생성형 AI의 출현과 현재 기술 환경, 유럽 인공지능법 발효 및 향후 AI 규정에 따른 새로운 규제 프레임워크 등이 반영되었다. 텔레포니카는 추가적으로 AI 시스템의 탄소배출 발자국 최소화와 관련된 "Green AI"에 대한 새로운 원칙을 추가할 계획이다. 텔레포니카의 「AI 원칙: AI Code of conduct」에 따르면, 현재 AI 행동 강령은 신뢰할 수 있고 사람 중심의 AI 채택을 촉진하고 건강, 보안, 기본 권리 및 환경에 대한 높은 수준의 보호를 보장하는 것을 목적으로 한다. 윤리원칙에서 규정하고 있는 사항들은 다음과 같다.

〈텔레포니카 AI 원칙: AI Code of conduct〉

- 인간중심의 AI
  - 우리는 AI가 인권을 존중하고 증진함으로써 세상을 더욱 인간답게 만드는 데 기여하기를 바랍니다.
  - 우리는 개인의 무결성을 보존하고 취약 계층을 보호하며 AI의 잠재적인 부정적 영향을 피하기 위해 최선을 다하고 있습니다.
  - 자율성, 인간의 존엄성, 선택의 자유를 지키기 위해서는 인간의 감독이 중요하다고 믿습니다.

- 투명하고 설명 가능한 AI
  - 사용자 신뢰를 높일 수 있도록 모델이 달성한 결과의 논리를 이해하려고 노력합니다. 성능과 설명 가능성 사이의 공정한 균형을 유지하기 위해 노력합니다.
  - 사람들이 AI와의 상호작용을 인지할 수 있도록 합니다.

- 공정하고 포용적인 AI
  - 우리는 공정하고 신뢰할 수 있는 결정을 내릴 수 있도록 AI 시스템이 제공하는 결과의 정확성을 높이고 있습니다.
  - 우리는 AI의 대표성과 접근성, 포용성, 공평성을 보장하고자 합니다.
  - 우리는 애플리케이션이 편견과 차별적인 영향을 미치지 않도록 노력합니다.

- 개인정보 보호와 보안을 존중하는 AI
  - 우리는 데이터 보호 및 개인정보 보호에 대한 권리를 존중할 것을 약속합니다. 또한 개인정보 보호 설계 방법론을 사용합니다.
  - 설계에 의한 보안 접근 방식에 따라 견고하고 강력한 AI 시스템을 보장하기 위해 노력합니다. 또한 추적성은 AI 시스템의 사이버 보안을 보장하는 데 필수적이라고 믿습니다.

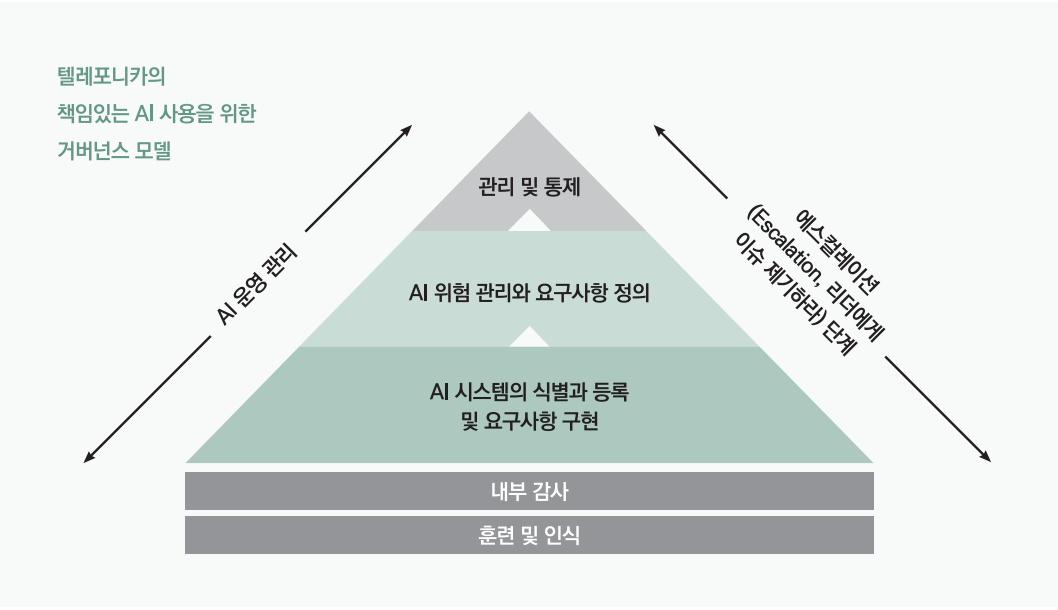
- 환경에 헌신하는 AI
  - 환경을 보존하고 지속가능성을 증진하며 기후 위기를 완화하기 위한 핵심적인 도구로서 AI를 장려합니다.
  - 환경 영향을 평가 및 최소화하고, 탄소 발자국을 줄이며, AI 시스템의 에너지 효율성을 최적화하기 위해 노력합니다.

- 가치사슬을 통한 책임과 책임성을 갖춘 AI
  - 책임과 의무는 AI로 대체할 수 없다고 굳게 믿고 있습니다.
  - 우리는 가치사슬에서 내린 의사결정의 추적성을 보장하기 위해 노력하며, 이는 직원 및 제3자와 협력할 때도 마찬가지입니다.
  - 역할과 책임을 정의하고 위험을 식별 및 완화할 수 있을 뿐만 아니라 AI 시스템의 감사 가능성을 보장하는 거버넌스 모델을 갖추고 있습니다.

출처: 텔레포니카(Telefonica), AI principle: AI code of conduct (2024.6.26)

한편, 텔레포니카는 2023년 12월부터 윤리원칙과 신뢰성 평가에 기반한 내부 AI 거버넌스 모델(규정)을 구축했다. 텔레포니카는 이 거버넌스 모델을 통해 책임 있는 기술에 대한 윤리적 약속과 참여하고 있는 국제 협력 프레임워크에 대응하고 있다. AI 거버넌스는 3대 기본사항(pillar)인 글로벌 가이드라인, 자체 규율, 규제 프레임워크를 바탕으로 구축되었다.

〈텔레포니카(Telefónica)의 AI 거버넌스 모델〉



출처: 텔레포니카(Telefónica) 홈페이지
https://www.telefonica.com/en/communication-room/blog/an-artificial-intelligence-governance-framework-at-telefonica/

텔레포니카가 AI 거버넌스를 구성한 과정을 살펴보면 다음과 같다. 먼저 2019년에 텔레포니카는 조직 내에 AI원칙을 구현하기 위해 6가지 주요 요소(①AI 원칙, ②인식 제고 및 교육 프로그램, ③다양한 AI 애플리케이션을 평가하기 위한 설문지, ④설명 가능성, ⑤공정성 및 환경 영향과 같은 기술 도구, ⑥거버넌스 모델)로 구성된 방법론을 설계하였다. 텔레포니카는 이 방법론을 이용해 직원을 대상으로 온라인 교육과정을 개설하여 AI의 기본사항 및 기술 사용 시 발생할 수 있는 기본적인 윤리적 질문에 대해 설명했다. 다음으로는 2020년부터 윤리적 원칙을 바탕으로 회사의 제품 및 서비스에 대한 윤리 평가 설문지를 제작했다. 2022년에는 설문 결과를 바탕으로 시범 AI 거버넌스 모델을 설계하여 파일럿을 운영하고 2023년 12월에 파일럿 경험을 바탕으로 텔레포니카의 공식 AI거버넌스 모델을 구축했다. 2024년에는 전사적으로 거버넌스 모델을 배포 및 시행하면서 점차 발전시켜가고 있다.

참고

- SPRI 소프트웨어정책연구소(장진철 외), 이슈리포트: 책임 있는 AI를 위한 기업의 노력과 시사점(2024.8.7)
- LG AI연구원 홈페이지 | https://www.lgresearch.ai/about/vision
- LG AI연구원, 2023년 LG AI 윤리책무성 보고서(2023)
- UNESCO, NEWS: "Insights from Practice: Telefónica's AI governance journey" https://www.unesco.org/en/articles/insights-practice-telefonicas-ai-governance-journey (2024.2.2/업데이트: 2024.8.5)
- Telefónica, "An Artificial Intelligence Governance Framework at Telefónica"(2023.12.6)
- Telefónica, Telefonica Artificial Intelligence Principles: AI Code of conduct (2024.6.26)
- Telefónica, Telefónica's Artificial Intelligence Principles| https://www.telefonica.com/en/wp-content/uploads/sites/5/2024/07/artificial-intelligence-principles-infographics.pdf

부패방지를 위한 생성형 AI

보고서리뷰

■ 보고서: OECD, Generative AI for Anti-corruption and Integrity in Government Taking Stock of Promise, Perils And Practice(2024.6)

(이 글을 읽는 데 약 8분이 소요됩니다.)

생성형 AI 기술은 우리의 일상에도 확산되고 있다. 최근에는 생성형 AI의 한 유형이자 거대 언어 모델(LLM, Large Language Model)[1]의 발전으로 그 영향력이 더 커지고 있다. 이러한 환경에서 정부는 이를 규제할 뿐만 아니라 빠르게 발전하는 기술을 활용하는 것도 필요하다. 부패방지 및 청렴성을 위해 노력하고 있는 '부패방지 입안 및 행정가(예: 부패방지 기관, 최고감사기관, 내부감사 및 통제 기능, 정부 차원의 청렴 부패방지 부처 등)' 역시 업무에 생성형 AI를 도입 및 활용할 수 있다.

이와 관련하여 OECD는 2024년에 'Generative AI for Anti-corruption and Integrity in Government Taking Stock of Promise, Perils And Practice' 보고서를 발간했다. 이 보고서에 따르면 부패방지 입안 및 행정가가 업무에 생성형 AI와 LLM을 활용하면 이점이 있는 반면 어려움이나 잠재적 위험도 존재한다. 이에 보고서는 AI가 부패방지 및 청렴활동에 활용되고 있는 현황과 사례를 제공하며 고려해볼 과제, 그리고 AI기술 활용에서 비롯되는 위험을 방지하는 방안 등을 소개한다.

보고서는 OECD 회원국 정부와 공공기관을 주요 대상으로 작성되었으나, 부패방지와 관련된 기관 및 행정가를 설문조사한 내용을 담고 있어, 기업의 청렴윤리경영 담당자가 부패방지에 AI를 활용함에 있어 일부 내용을 참고할 수 있을 것으로 판단된다. 이번 보고서 리뷰에서는 관련 내용을 발췌 요약하여 소개하고자 한다.

1) 대량의 텍스트 데이터를 학습하여 자연어 처리 작업을 수행하는 인공지능 모델로 문의나 명령을 해석하고 인간과 유사한 언어로 응답을 생성하는 기계 학습 알고리즘

부패방지 및 청렴성 향상을 위한 기회와 과제

부패방지 분야에서 응용되는 LLM

생성형 AI, LLM의 등장은 부패방지 입안 및 행정가에게 업무 효율성과 영향력을 향상시키는 새로운 방법이다. 부패방지 입안 및 행정가들은 생성형 AI 중에서도 LLM에 주목한다. LLM은 주로 잠재적 위험 쿼리(Query)[2]등을 이용해 문서나 데이터 소스로부터 자동으로 사기를 탐지하거나 대량의 문서와 데이터를 처리를 도움으로써 부패방지 및 조사기관의 업무 수행을 지원할 수 있다. 예를 들면, LLM은 대량의 텍스트를 정리하여 우선순위 결정, 근본 원인 분석, 패턴 인식 등을 돕는다. 감사·조사관은 이러한 기술을 통해 수고와 오류를 줄이고 아직 생성형 AI가 대체하지 못해 인간의 판단과 전문성이 필요한 고부가가치 업무에 더욱 집중할 수 있다. 또한 LLM은 공공 조달 담당자가 기업 또는 잠재적인 계약 업체에 대한 대량의 데이터를 분석해서 사기나 부패 위험을 식별하는 데에도 사용할 수 있어 공공 지출 영역에서의 청렴성도 증진시킬 수 있다.

내부 운영 개선 및 부패·사기 방지 활동 강화

조직의 내부 운영과 반부패 및 사기방지 활동 영역은 서로 연관되어 있다. 때문에 한 영역에서 생성형 AI나 LLM을 사용하여 정보를 처리·분석한 성과는 다른 영역의 효율성을 향상시킬 수 있다. 보고서에 따르면 생성형 AI, LLM를 활용하면 조사 및 감사 프로세스의 증거 수집과 문서 검토 분야와 부패 및 사기 방지활동에서 문서 분석과 패턴 인식에서 가장 큰 이점을 얻을 수 있는 것으로 나타났다.

LLM 구현의 어려움과 조언

부패방지 입안 및 행정가들의 생성형 AI, LLM를 활용에 대해 가장 우려되는 사항은 기술 및 경험 부족과 데이터 프라이버시 및 보안을 꼽았다. 그 다음으로는 예산 제약, 데이터 품질, IT 규제, AI가 생성한 결과에 대한 우려(예: 편견과 환각)도 과제로 꼽혔다.

생성형 AI 및 LLM을 시범적으로 운영할 때에는 문서 요약 및 보도자료 작성, 사용자 질문에 대한 답변 등 비교적 위험도가 낮은 프로세스부터 통합하는 것을 추천한다. 생성형 AI 적용 범위를 확대하기 전에 이용자의 역량을 구축할 수 있으면서 위험도가 낮아 재정적으로나 규정 준수 실패로 인한 비용이 덜 들 수 있기 때문이다. 또한 생성형 AI의 적용범위를 확대하기 위해 '컴퓨터 성능, 데이터 스토리지, 소프트웨어' 등 IT인프라 자원을 마련하는 것이 필요하다.

---

2) 데이터 베이스에서 속성 데이터를 조사하는 것

핵심 과제

현재 생성형 AI, LMM을 사용하는 데 있어 개선 노력이 필요한 과제들은 다음과 같다.

언어 장벽 극복	• 대부분의 LLM 모델은 영어로 학습되어 있기 때문에 이를 배포하고 사용하는데 한계가 있다. 이를 극복하기 위해 일부 국가에서는 현지 언어로 학습된 LLM개발에 투자한다.
안전장치 필요성	• 생성형 AI 개발 및 구현과 관련된 과제 중에서 규정 준수 및 윤리 문제는 가장 낮은 순위로 꼽혔다. 그러나 AI가 고급업무에 사용되고, AI규제가 발전할수록 규정 준수에 대한 압력을 받을 것이다.
편견과 차별 완화	• 발생 원인으로는 모델 기초 데이터가 되는 샘플의 모집단에 대한 대표성 부족, 알고리즘 학습에서 지속적인 오류 결과 값 도출(통계적 편향), 개인이 제공하는 텍스트의 불균형한 학습 등이 있다. • 편견과 차별을 완화하기 위해 편견의 위험과 결과 모두를 인식해야 한다. 그러나 이에 대해 전반적인 관심이 부족하고 완화 조치가 이루어지는지 확인하기 위한 책임과 투명성 메커니즘이 부족하다.
설명 가능성 확보	• LLM를 평가하고 설명하기는 어려울 수 있다. 그러나 AI 모델의 결과물의 해석가능성 및 설명가능성 등을 고려하여 LLM모델을 모니터링하고 평가하는 것은 중요한 고려사항이다.

위험 완화

생성형 AI는 부패방지 및 청렴 업무를 수행하는데 위험을 초래하기도 한다. 위험의 예로는 적대적 공격 (Adversarial Attack)[3]과 피싱, 모방(사칭), 도용 등이 있으며, 딥페이크의 확산으로도 위험성이 더욱 커지고 있다. LLM 사용으로 부패 감지가 어려워지거나 공직자가 부정을 저지르기 쉬워지기도 한다. 또한 자동화된 의사결정에 지나치게 의존하면 대중의 신뢰를 잃거나 윤리적 문제를 야기할 수 있다. 생성형 AI는 고의 또는 의도 없이 허위정보를 확산할 위험이 있으며, 거짓 또는 반복되는 댓글, 스팸 처리 등으로 정책 결정과정을 방해할 위험도 존재한다. 이와 관련하여 호주 뉴사우스웨일즈주 독립부패방지위원회(ICAC)는 부패방지 업무에 생성형 AI를 도입할 때 나타날 수 있는 잠재적 기회와 위험을 제시한다. 이 내용을 정리하면 아래의 표 같다. ICAC는 특히, 잠재적 위협 중 실제로 나타나는 문제인 3번과 5번을 강조한다.

3) 인공지능 모델의 취약점을 공격하는 머신러닝(Machine Learning) 공격기법

〈부패방지 업무에 대한 AI의 잠재적 위험〉

잠재적 기회	잠재적 위험
1. 대규모 데이터 세트의 필터링 2. 분류 및 분석을 통해 지능을 향상시키는 AI의 능력 3. 패턴 인식 4. 예측 및 모델링 5. 감정 분석 6. 데이터의 이상 징후 감지 7. 데이터 통합 및 다중 소스 분석 8. 인간의 재량권 제한을 통한 부패 기회 감소	1. AI의 딥페이크 제작 능력 2. AI를 이용한 사이버 범죄 강화 3. 공무원의 AI 악용 4. AI에 대한 의존 5. AI를 사용한 정부 문서 위조 6. 민주주의와 공적 담론에 대한 위협 7. 아웃소싱과 관련된 위험

출처: OECD, Generative AI for Anti-corruption and Integrity in Government Taking Stock of Promise, Perils And Practice(2024)

보고서는 이처럼 LLM이 부패방지 업무나 외무환경에 영향을 미칠 수 있는 새로운 위험을 야기하거나 강화할 수 있지만 그 외의 다른 위험은 본질적으로 내부적인 것이므로 통제 및 위험 완화 조치를 개발할 때 이를 고려해 볼 만하다고 제시한다.

참고

• OECD, Generative AI for Anti-corruption and Integrity in Government Taking Stock of Promise, Perils and Practice (2024.6)
• 방송통신위원회&NIA, 생성형 AI 윤리 가이드북(2023)

행동하는
윤리경영

AI 시스템 개발 및 도입 시 점검사항

(이 글을 읽는 데 약 9분이 소요됩니다.)

글로벌 경영 전략 컨설팅 회사인 액센츄어(Accenture)와 스탠퍼드 대학교 인간중심 인공지능 연구소(HAI)가 실시한 조사 결과에 따르면 유럽, 북미, 아시아의 기업은 AI 도입 전략에 '책임 있는 AI의 요인(공정성, 투명성, 설명 가능성, 개인정보 및 데이터 거버넌스, 신뢰성 및 보안)' 중 하나 이상을 고려하고 있다. 기업이 위험 완화 를 위해 기업전략에 고려하는 조치로는 '데이터 거버넌스'의 비율이 높게 나타난 반면, '투명성 및 공정성'의 비율은 상대적으로 낮은 경향이 있어 기업의 인식과 조치 향상이 요구된다.

또한 최근 유럽연합의 EU 인공지능법(EU AI Act), 미국 백악관의 '안전하고 보안적이며 신뢰할 수 있는 인공지능(AI) 행정명령'등 위험성이 높은 AI에 대한 규제 조치들이 진행되고 있으므로 기업의 고위험 AI의 안전을 확보를 위해 신속히 프로세스를 마련하고 확산시켜야 한다.

이에 이번 행동하는 윤리경영에서는 AI 윤리 및 안전을 위해 과학기술정보통신부와 한국정보통신기술 협회가 제작한 가이드 라인인 '2024 신뢰할 수 있는 인공지능 개발 안내서'(2023) 등을 참고하여 AI 시스템을 계획하거나 설계 또는 도입할 때 특히, AI 거버넌스와 관련하여 청렴윤리경영 담당자가 고려해볼 수 있는 사항들을 알아보고자 한다.

1 AI 시스템 개발 및 도입 시 점검할 질문들

각국 정부는 AI를 규제하는 정책과 입법을 서두르고 있다. 규제 강화는 AI 산업의 신뢰성을 높이고 장기적으로는 글로벌 시장에서 경쟁력을 강화하는 데 기여할 수 있다. 또한 EU 인공지능법 등은 글로벌 AI의 기준이 될 가능성이 크므로 규제를 준수하는 기업은 글로벌 시장에서도 경쟁 우위를 확보할 수 있을 것으로 보인다. 따라서 기업이 AI 기술의 사용을 고려한다면 AI위험을 이해하고 이에 대한 식별 및 관리 조치를 취해야 한다.

'2024 신뢰할 수 있는 인공지능 개발 안내서' 보고서는 미국, 유럽 등 주요 선진국과 국제기구에서 발표한 권고안 및 가이드를 바탕으로 기업이 AI를 사용하는데 있어, 최소한의 신뢰성을 확보하는 방안과 주요 사항을 이해하고 자율 점검할 수 있는 내용을 제공하고 있다. 보고서에 따르면 위해 비즈니스 결정권자는 계획 및 설계단계의 주요 행위자다. 일반적으로 비즈니스 결정권자는 기획자 등과 협력하여 시스템 전체 생명주기에 걸친 신뢰성을 확보하고 요구사항을 검토하고 적용방안을 마련하는 것이 요구된다. 청렴윤리경영 담당자는 AI 시스템의 위험요소를 관리하고 관리체계를 구축하는데 다음과 같은 네 가지 사항들을 참고해 볼 수 있다.

첫째로, AI 시스템에 대한 위험관리를 계획하고 및 수행한다. 시스템이 구현 및 운영되는 과정에서 발생할 수 있는 모델 오인식, 기능 오동작, 보안 및 개인정보 이슈 등의 위험 요소를 사전에 인식한다. 그 뒤 위험의 심각성 및 파급효과를 분석하여 대응 방안을 마련해야 한다.

둘째로, 먼저, AI 관련 법, 규제, 정책, 표준 및 지침을 정리하여 내부적으로 준수할 규정을 수립한다. 또한 이를 관리·감독하는 AI 거버넌스 체계를 구성해야 한다(예: AI 시스템 생명주기에 따른 조직의 업무 역할과 책임을 명문화). AI 시스템의 사회적인 영향과 결과를 예측하고 그에 대비하는 조직을 구성하는 것은 신뢰성 확보에 중요한 요소이다.

셋째로, AI 시스템의 신뢰성 테스트 계획을 수립해야 한다. AI는 추론 연산과정이 예측 불가능하다는 불확실성(uncertainty)을 내포하고 있다. 안전성과 신뢰성을 확보하려면 이를 줄이는 것이 중요하다. 따라서 소프트웨어의 품질 확인 테스트 외에도 추가적인 테스트를 통해 AI 시스템의 신뢰성을 확인할 필요가 있다. 테스트를 계획할 때는 AI 시스템의 복잡도(complexity)와 운영환경을 고려하고, 생명주기 전 단계에서 계획에 따라 정기적·지속적으로 실시해야 한다.

마지막으로, AI 시스템의 추적가능성과 변경이력을 확보해야 한다. 문제 원인을 추적할 수 있도록 AI 시스템 운영 단계에서 시스템 로그, 데이터 모니터링, AI 모델과 사람 간 의사결정 기여도 추적, 변경이력 관리 등의 방안을 확보한다. 이러한 사항들은 아래의 질문들을 통해 검토해 볼 수 있다.

〈신뢰할 수 있는 AI 개발을 위한 체크리스트-비즈니스 결정권자의 검토사항〉

생명주기	요구사항 및 체크리스트	Yes·No·N/A
생명주기	**요구사항 01 : AI 시스템에 대한 위험관리 계획 및 수행**	☐ ☐ ☐
	AI 시스템의 생명주기에 걸쳐 나타날 수 있는 위험 요소를 도출하고 파급효과를 파악했는가?	☐ ☐ ☐
	AI 기술 적용을 어렵게 만드는 위험 요소가 있는지 확인하였는가?	☐ ☐ ☐
	위험 요소별 완화 또는 제거 방안을 마련하였는가?	☐ ☐ ☐
	위험 요소의 파급효과가 감소하였는지 확인하였는가?	☐ ☐ ☐
	요구사항 02 : AI 거버넌스 체계 구성	☐ ☐ ☐
	내부적으로 준수해야 할 AI 거버넌스에 대한 지침 및 규정을 마련하였는가?	☐ ☐ ☐
	AI 거버넌스를 위한 조직을 구성하였는가?	☐ ☐ ☐
	AI 거버넌스를 위한 조직은 충분히 훈련된 인력으로 구성하였는가?	☐ ☐ ☐
	AI 거버넌스에 대한 내부 지침 및 규정 준수 여부를 감독하고 있는가?	☐ ☐ ☐
	AI 거버넌스 조직이 신규 및 기존 시스템의 차이점을 분석하였는가?	☐ ☐ ☐
	이용 빈도가 낮은 타 시스템의 개선 및 통폐합을 통해 구현 가능한지 분석하였는가?	☐ ☐ ☐
	요구사항 03 : AI 시스템의 신뢰성 테스트 계획 수립	☐ ☐ ☐
	테스트 환경 결정 및 설계 시 AI 시스템의 운영환경을 고려하였는가?	☐ ☐ ☐
	가상테스트 환경이 필요한 AI 시스템의 경우, 시뮬레이터를 확보하였는가?	☐ ☐ ☐
	AI 시스템의 테스트 설계에 필요한 협의 체계를 구성하였는가?	☐ ☐ ☐
	- AI 시스템의 기대 출력을 결정하기 위한 협의 체계를 구성하였는가?	☐ ☐ ☐
	- 설명가능성 및 해석가능성 확인을 위한 사용자 평가단을 구성하였는가?	☐ ☐ ☐
	요구사항 04 : AI 시스템의 추적가능성 및 변경이력 확보	☐ ☐ ☐
	AI 시스템의 의사결정에 대한 추적 방안을 수립하였는가?	☐ ☐ ☐
	- AI 시스템의 의사결정에 대한 기여도 추적 방안은 확보하였는가?	☐ ☐ ☐
	- AI 시스템의 의사결정 추적을 위한 로그 수집 기능을 구현하였는가?	☐ ☐ ☐
	- 지속적인 사용자 경험 모니터링을 위해 사용자 로그를 수집 및 관리하고 있는가?	☐ ☐ ☐
	학습 데이터의 변경 이력을 확보하고, 데이터 변경이 미치는 영향을 관리하였는가?	☐ ☐ ☐
	- 데이터 흐름 및 계보(lineage)를 추적하기 위한 조치를 마련하였는가?	☐ ☐ ☐
	- 데이터 소스 변경에 대한 모니터링 방안을 확보하였는가?	☐ ☐ ☐
	- 데이터 변경 시, 버전관리를 수행하였는가?	☐ ☐ ☐
	- 데이터 변경 시, 이해관계자를 위한 정보를 제공하는가?	☐ ☐ ☐
	- 신규 데이터 확보 시, AI 모델의 성능평가를 재수행하였는가?	☐ ☐ ☐

출처: 과학기술정보통신부·한국정보통신기술 협회, 2024 신뢰할 수 있는 인공지능 개발 안내서 – 일반 분야(2024.2)]

한편, AI 시스템을 개발하지 않고 조달하는 기업의 경우, AI 거버넌스 관리를 위해 공급업체가 신뢰할 수 있는 AI를 제공하고 있는지 점검해야 한다. 세계경제포럼(WEF)가 발간한 보고서 'Adopting AI Responsibly: Guidelines for Procurement of AI Solutions by the Private Sector'(2023.6)는 사용자의 중요한 의사결정에 AI 시스템이 예측이나 권장을 통해 영향을 미칠 것이므로 AI가 윤리적이며 책임감 있고 신뢰할 수 있도록 운영되는 것이 필수적이라고 언급한다. 그러려면 기업은 AI를 도입할 때 (1)비즈니스 전략과 일치 여부 (2)사업 사례 (3)기술 및 데이터 통합 (4)AI 시스템과 윤리적 정렬 (5)위험평가 (6)민첩하고 경제적인 시스템 등을 확인해야 한다. 이와 함께 해당 보고서는 AI 시스템을 조달하기 위한 프로세스 각 단계에서 공급업체를 평가하고 이해하는 데 사용하는 질문을 제공한다. 이 질문들 중 AI 윤리와 거버넌스에 대한 내용을 정리한다면 아래의 표와 같다. 조달하는 업체뿐만 아니라 공급업체도 신뢰있는 AI 거버넌스를 관리하고 AI 시스템을 제공하기 위해 참고해 볼 수 있을 것이다.

〈조달시 AI 거버넌스, 위험 및 컴플라이언스 관리: 질문사항〉

설명
1. 이 모델의 데이터 수집 대상은 어떤 집단인가? (AI 모델링은 데이터에 기반하기 때문에 개인정보보호 및 동의, 개인과 관련된 규정 측면에서 데이터 사용이 대상에 미치는 영향을 고려하는 것이 필수적이다.)
2. 공급업체는 사이버 보안 위험을 어떻게 관리하고 있는가?
3. 공급업체는 다른 지역(물리적)에서 운영할 때 발생할 수 있는 지정학적 위험을 검토했는가?
4. 프로젝트와 관련된 위험이 명확하게 정의되어 있는가?
5. 공급업체가 AI 모델 구축과 관련된 규칙 및 규정을 준수하고 있는가?
6. 공급업체는 감사 및 규정 준수 요건에 어떻게 대비하는가?
7. 공급업체가 모델에서 국제 표준 및 인증을 구현하고 있는가?
8. 공급업체는 어떤 조직 관행을 권장하는가?
9. 계약서에 규정 준수와 관련된 모든 요소가 포함되어 있는가?

출처: WEF, Adopting AI Responsibly: Guidelines for Procurement of AI Solutions by the Private Sector(2023.6)

참고

- 과학기술정보통신부·한국정보통신기술 협회, 《2024 신뢰할 수 있는 인공지능 개발 안내서 – 일반 분야》(2024.2)
- SPRI 소프트웨어정책연구소(장진철 외), 이슈리포트: 책임 있는 AI를 위한 기업의 노력과 시사점(2024.8.7)
- 한국지식재산연구원, "미국 백악관, AI 행정명령 시행 270일 이후 이행 점검 및 성과 발표"(2024-08-06) https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=US&po_no=23031
- WEF, Adopting AI Responsibly: Guidelines for Procurement of AI Solutions by the Private Sector(2023.6)
- 테크월드뉴스, "유럽서 올해 'AI 규제 법안' 최초 통과…해외 진출 국내 기업 대비책 필요"(2024.8.22) | https://www.epnc.co.kr/news/articleView.html?idxno=305399
- ZDNET korea, "AI은행원 등 생성형 AI 주시하는 금융권…거버넌스 미리 구축해야"(2024.2.11) https://zdnet.co.kr/view/?no=20240208082515

문화 속
기업윤리

과학적으로 다루는 기업부패,
도서 '언더그라운드 이코노미'

*이미지 출처: 예스24

AI 기술은 대규모의 데이터를 처리하고 분류하는 데 상당한 효율성을 제공한다. 한 예로 정부의 각종 프로세스를 자동화하여 부패가 개입될 수 있는 부분을 사전에 차단할 수도 있는 것처럼 딥러닝(deep-learning) 기술은 부패 범죄의 패턴을 학습하고 미리 탐지할 수 있다. 반면에 생성형 AI와 같은 최신 기술이 범죄에 활용되는 것을 보면 긍정적인 효과만 있지는 않을 것으로 예상된다.

도서 '언더그라운드 이코노미: 부패의 메커니즘'은 부패와 관련된 다양한 질문에 대해서 사례와 이론을 들어 답을 제시하고 있다. 예를 들어 부패 전략을 사용하는 기업이 성공할 수 있을까라는 질문에 대해 부패 활동은 기업의 지배구조를 비효율적으로 만들고, 사법리스크 및 여러 부대 비용으로 기업 성과에 부정적인 영향을 미칠 것이라고 생각되며 반대로, 부패한 기업이 희소한 자원에 접근할 수 있는 기회가 많고, 까다로운 규제장벽을 피할 수 있으므로 성과가 좋을 수 있다는 주장할 수도 있다. 저자는 실증분석 사례를 들어 부패 행위는 기업성과에 그 자체로 직접적으로 방해나 도움이 된다고 보기는 어려우며, 과학적이고 계량적인 접근을 통해 합리적 판단이 필요하다고 설명한다. 마찬가지로 이번 호에서 다룬 AI와 같은 기술 발전은 부패를 막는데 다양하게 활용될 수도 있는 반면 국가와 일부 거대기업이 그것을 오용하고 독점하여 부패상황을 악화시킬 위험성도 유념해야 함을 보여준다.

부패를 효과적으로 줄이기 위해서는 부패를 반드시 근절해야겠다는 최고 지도자(CEO) 및 국민(구성원)의 각성, 부패를 실제로 통제할 수 있는 관료들의 우수성 등 다양한 요인이 뒷받침되어야 한다. 저자는 이 뿐만 아니라 지하(언더그라운드)에서 다루어지는 부패를 공론화하고, 좀더 과학적(경제학적이고 통계기법을 활용한)인 방법으로 부패현상을 발견하고 탐지하는 노력이 필요함을 도서를 통해 강조하고 있다.

국민권익위원회 청렴윤리경영 동향

뉴스클립

'채용기준 부재'한 공공기관에 일원화된 공정채용 기준 마련된다

국민권익위원회는 '2024년 청년주간(9월 21일~27일)'을 맞아 '기타공직유관단체 공정채용 운영기준'을 마련하고 398개소의 기타공직유관단체를 대상으로 자체 규정화할 것을 권고했다. 현재 공사 · 공단 등 공직유관단체 상당수(1,031개소)는 유형에 따라 「공공기관의 운영에 관한 법률」 등 상위 법령과 지침 등에 따른 표준화된 채용절차를 운영하고 있다. 그러나 관련 법령 · 지침을 적용받지 않는 기타공직유관단체는 연간 채용규모가 약 9,900명에 달함에도 불구하고 별도의 절차와 기준이 미비해 잠재적인 불공정 채용 위험성이 제기되어 왔다. 이번 국민권익위 권고는 「공공기관의 운영에 관한 법률」 등 관련 법령 · 지침과 다수 공공기관에서 널리 시행하고 있는 공정채용 절차와 기존 권고사항 등을 종합 일원화한 것이다. 채용 과정에서 ▲임직원 가족 우대채용 금지 ▲원서접수 시 불필요한 인적사항 요구 금지 ▲블라인드 채용원서 활용 ▲심사위원 위촉 시 외부 위원 포함 ▲채용비리 피해자 구제 등 채용 절차의 공정성을 높이는 내용으로, 현장에서 널리 활용될 수 있도록 채용계획 수립부터, 공고, 심사전형 및 합격자 결정에 이르기까지 채용 전 단계에 걸쳐 준수하여야 할 37개 항목으로 구성되었다. 또한 제도개선 권고 이후에는 기관별 자체 규정화 이행 여부를 점검하는 등 체계적인 관리·감독을 병행 추진할 예정이다.

▪ 국민권익위원회 2024년 9월 24일
https://www.acrc.go.kr/board.es?mid=a10402010000&bid=4A&act=view&list_no=78706

국민권익위, 중남미 스페인어권 국가 대상 반부패 연수 시행

국민권익위원회 청렴연수원은 9월 3일부터 9월 11일까지 '다국가 반부패 역량강화 연수' 스페인어 과정을 운영했다. 콜롬비아 대통령실, 페루 총리실, 볼리비아 법무투명성부, 파라과이 감사원 등 4개국 반부패 관계기관 공무원 총 15명이 참여했다. 연수 프로그램은 2012년 유엔 공공행정상 대상을 수상한 공공기관 종합청렴도 평가를 비롯해서 ▲부패영향평가 ▲부패·공익 신고자 보호제도 등 국제적으로 인정받고 있는 대한민국의 우수 반부패 제도를 중심으로 구성되었다. 또한 '청렴포털', 온라인 공직자 재산신고 시스템, '나라장터' 등 부패 예방에 효과적인 한국의 주요 전자정부 시스템을 소개하고, 개도국들의 부패 취약분야인 공공계약과 조달 분야에서 부패를 적발하는 감사기법도 공유했다. 우리나라는 경제 발전과 부패문제 해결을 동시에 성공한 모범 사례로서, 청렴포털과 나라장터 등 디지털 기술을 활용한 최첨단 반부패 플랫폼을 통해 다양한 행정 분야를 아우르는 실질적인 반부패 정책을 집행하고 있다. 이와 같은 우리나라의 반부패 정책에 대해 경제 개발과 함께 반부패를 정착시키고자 하는 개발도상국에서 특히 관심을 갖고 있다. 국민권익위는 이러한 교육 수요에 따라 영어 연수 과정 외에 중앙아시아와 동유럽 국가들을 위한 러시아어 과정(2020년), 아프리카 국가를 대상으로 한 불어 과정(2023)을 신설하고 올해는 중남미 국가들을 대상의 스페인어 연수 과정도 개설하여 K-청렴을 다양한 언어로 전 세계에 알릴 수 있게 됐다.

▪ 국민권익위원회 2024년 9월 4일
https://www.acrc.go.kr/board.es?mid=a10402010000&bid=4A&act=view&list_no=77724

국내외 동향

EU 인공지능법(AI Act) 준수협약에 115개사 참여

유럽연합(EU) 인공지능법(AI Act) 준수협약에 삼성과 구글, 마이크로소프트(MS) 등 115개사가 참여했다. EU는 AI 기술이 사람이나 사회에 해를 끼칠 우려를 고려해 AI 기업이 피해를 막기 위해 의무적으로 기술에 담아야 할 내용을 규정한 인공지능법을 마련해 올 8월 발효했다. 세계 최초의 포괄적 AI 규제로 평가되는 법이다. 고위험 AI규제 등 세부 규정은 전면 발효되는 2026년 8월까지 차례로 발효할 예정이다. EU 내에서 제품·서비스를 제공하는 EU AI법을 준수하지 않을 경우 전 세계 연 매출의 1.5~7%의 과징금을 부과받을 수 있다. EU는 전면 발효까지 2년이 남은 만큼 IT기업에 AI법 준수협약에 참여를 독려해 왔다. 아직 법적 구속력은 없지만 기업들이 자발적으로 참여해 법 제정 효과가 나타나는 시점을 최대한 앞당기겠다는 취지다. 삼성을 비롯한 115개 서약 참여 기업은 고위험 AI 기술로 분류될 만한 자사 시스템을 사전 점검하고, 법 준수를 위한 조직 내 AI 거버넌스 전략도 수립할 예정이다. 참여 기업 중 절반 이상이 AI 기술 사용 때의 인적 감독을 보장하고 딥페이크 등 특정 유형의 AI 기반 콘텐츠에 별도 표기를 하는 등 추가적인 노력을 하기로 했다고 EU 측은 전했다.

▪ 이데일리 2024년 9월 25일 https://n.news.naver.com/article/018/0005843465?sid=105

일본 ESG 투자 붐...ESG 채권 발행 2년 새 두배 급증

일본에서 ESG 투자에 대한 관심이 뜨겁다. 이러한 현상의 이유로 ESG에 대한 정부의 관심이 이 분야에 대한 기업과 금융시장의 관심을 불러일으킨 가운데 투자자들 사이에서 ESG 투자가 장기적으로 더 지속가능한 투자라는 인식이 자리를 잡고 있다는 점이 거론된다.

일본증권업협회(JSDAQ)에 따르면 일본에서는 2021년과 비교했을 때 ESG 채권 발행 규모가 2023년 6조7000억엔(약 61.4조원)으로 두 배로 증가했다. 전문가들은 이렇게 된 데는 일본이 2020년에 2050년까지 온실가스 순배출 제로를 달성하겠다고 한 공약이 긍정적인 영향을 미친 것으로 분석했다. 전문가들은 이렇게 된 데는 일본이 2020년에 2050년까지 온실가스 순배출 제로를 달성하겠다고 한 공약이 긍정적인 영향을 미친 것으로 분석했다. 일본 정부가 세계 최초로 국채형 기후전환채권을 발행해 녹색 투자를 촉진하기로 한 것도 ESG 채권 시장 호황에 긍정적인 영향을 미친 것으로 분석됐다. 일본 재무성은 올해 2월에 약 7조엔(약 64조원) 규모의 10년 만기 기후전환채권을 발행했다. 일본은 2050년까지 온실가스 순배출 제로 달성을 위해 향후 10년간 20조엔(약 183조원)의 기후 전환채권을 발행할 예정이다. 채권 발행을 통해 얻은 자금은 일본의 순배출 제로 전환을 가속화하기 위한 다양한 프로젝트와 기술에 쓰겠다는 계획이다.

▪ 한국경제 2024년 9월 5일 https://www.hankyung.com/article/202408264866i
▪ ESG경제 2024년 8월 21일 https://www.esgeconomy.com/news/articleView.html?idxno=7622

IPEF 청정경제·공정경제협정 10월 발효

인도·태평양 경제프레임워크(IPEF)의 주요 구성 축인 청정경제·공정경제 협정이 오는 10월 순차적으로 발효된다. 산업통상자원부는 9월 24일 화상으로 열린 IPEF 장관회의에서 회원국 장관들이 청정경제·공정경제 협정이 각각 내달 11일과 12일 발효된다고 확인하면서 환영의 뜻을 나타냈다고 밝혔다. IPEF는 2022년 5월 출범하여 미국과 한국, 일본, 호주, 인도, 태국, 말레이시아, 인도네시아, 베트남, 필리핀, 싱가포르, 브루나이, 뉴질랜드, 피지 등 14개국이 참여하고 있다. IPEF는 무역(필러1), 공급망(필러2), 청정경제(필러3), 공정경제(필러4) 4개 협정으로 나뉜다. 앞서 쟁점이 많은 무역 협정을 제외한 공급망·청정경제·공정경제 협정이 우선 타결된 상태로, 이 중 공급망 협정은 지난 4월 가장 먼저 발효됐다. 새롭게 발효될 공정경제 협정은 부패 신고자 보호 강화, 정부 조달 과정에서 불법 행위 처벌 규정 도입 등 부패 방지와 조세 행정의 투명성 및 효율성을 제고하는 내용을 담고 있다. 청정경제 협정에는 참여국들이 청정에너지원을 포함한 에너지 생산 과정에서부터 탄소 저감 기술, 탄소 거래 시장에 이르는 에너지 산업 전 단계에서 기술, 규범, 표준 협력을 강화하는 내용이 담겼다. 산업통상자원부는 "정부는 청정경제·공정경제 협정의 국내 발효를 차질 없이 진행하고, IPEF의 가시적 성과 도출을 위해 인도·태평양 주요국과 협력을 강화하고 우리 기업들이 실질적 혜택을 받을 수 있도록 총력을 다할 계획"이라고 말했다.

▪ 연합뉴스 2024년 9월 24일 https://n.news.naver.com/article/001/0014943656?sid=101

미·EU '보호무역주의 확산' 속 국내 ESG 관심 필요

최근 미국의 생물보안법 등으로 중국 바이오기업에 대한 제재 등이 예고되는 가운데, 이에 대한 반사이익을 얻기 위해서는 ESG 요건 등에도 관심을 기울여야 한다는 주장이 제기됐다. 한화투자증권 리서치센터는 9월 12일 한국제약바이오협회의 KPBMA FOCUS에 '보호무역주의 확산과 ESG 대응과제'를 통해 이같이 밝혔다. 미국, EU 등 주요국은 첨단기술을 중심으로 자국 우선주의 정책을 추진하면서 중국을 견제하기 위한 수단으로 ESG, 특히 환경과 인권 이슈를 활용하고 있으며, 그 첨단기술은 반도체에서 AI, 바이오산업으로 범위가 확대되고 있다고 소개했다. 이에 미국, 유럽 등에 위치한 빅파마(Big pharma, 거대제약회사)는 거래 협력사를 대상으로 바이오의약품 생산 전 과정에 대해 탄소 중립 등 목표 달성을 촉진하고, 인권, 플라스틱 규제, 생물다양성 등 ESG 활동을 요구하고 있다. 이는 결국 기업 간 거래 조건으로 귀결되어 공급망 관리 중요성이 대두되고 있으며, ESG 활동은 선택이 아닌 필수라는 분석이다. 그러나 대한상공회의소가 국내 수출기업 205개사를 대상으로 '국내 수출기업의 ESG 규제 대응현황 정책과제'를 조사한 결과에 따르면, 국내 기업들의 ESG 수출규제 인식 및 대응 수준이 크게 미흡하다는 조사 결과가 나타난 바 있다. 한화투자증권 리서치센터는 "미국·EU 및 중국 간 경쟁이 심화될 경우, 국내 제약바이오 기업이 반사이익을 얻을 수 있으나 공급망에서 요구하는 ESG 요건을 충족하지 못하면 빅파마와 거래에서 배제될 수 있기 때문에 대응 방안을 마련해야 한다"고 지적했다.

▪ 메디컬타임즈 2024년 9월 12일 https://www.medicaltimes.com/Main/News/NewsView.html?ID=1160475&ref=naverpc
▪ 중앙일보 헬스미디어 2024년 9월 12일 https://jhealthmedia.joins.com/article/article_view.asp?pno=28697

생성형 AI 활용 윤리

웹툰

*참고: 방송통신위원회&NIA, 생성형 AI 윤리 가이드북(2023)

행사소식

New EU ESG Regulations and Impact on Korean Companies Doing Business in EU

EU의 규제 변화에 대응하기 위해 법무법인 율촌(Yulchon LLC)이 한국 법인에 직간접적으로 영향을 미칠 수 있는 주요 규정에 대한 실질적인 통찰력을 제공하는 세미나.

주최 법무법인 율촌
일정 2024년 10월 15일(화)
장소 법무법인(유) 율촌 39층 Lecture Hall
참고 https://www.yulchon.com/ko/seminars/seminars-view/596/page.do

24년도 지속가능경영확산사업 ESG 네트워크포럼

산업자원부 [지속가능경영확산사업]의 일환으로 대기업·중견·중소기업을 대상으로 분야별 다양한 이해관계자와 함께 K- ESG 상호 협력체계를 구축하여, 최신 국내외 ESG 이슈, 산업계 대응전략, 정책방향을 공유하고 소통하고자 개최되는 포럼.

주최 산업통상자원부
일정 2024년 10월 16(수)~11월 28일(목)
장소 오프라인(개별 확인)
참고 https://docs.google.com/forms/d/e/1FAIpQLSfvsbanSuZMN_SZ-EfIP7oSXVLYhd8HXLISNmYdM5um0QU74Q/viewform

독자 의견

청렴윤리경영 브리프스 10월호 내용이 도움이 되셨나요?

독자분들의 의견을 반영해 더 나은 콘텐츠로 개선하려 합니다.
아래의 질문에 대한 응답을 남겨주시면 5명을 추첨하여 기프티콘을 보내드립니다.

① 이번호에서 다룬 주제 또는 내용이 도움이 되셨나요?
② 어떤 부분이 마음에 드셨나요(또는 어떤 부분이 마음에 들지 않으셨나요)?
③ 기타 의견(다뤘으면 하는 주제, 기업 사례 등)

2024년 10월 20일(금)까지
(1) '의견남기기' 페이지(https://quiz.assist.ac.kr)에서 응답하시거나
(2) 국민권익위원회 민간협력담당관실(korea@assist.ac.kr) 앞으로
의견과 성함, 연락처(휴대폰 번호)를 보내주세요.
* 수집된 개인정보는 상품 발송을 위한 정보로만 활용되며, 추첨 이후 파기됩니다.

귀중한 의견에 감사드립니다!
기프티콘 추첨에 선정되신 분들은

김○지님, **안○규**님, **안○현**님, **황○연**님입니다.

민원상담 신청부터 결과확인까지 온라인으로 한번에 끝내기

법령·제도·행정 등 민원상담은 ▶ 국민신문고 www.epeople.go.kr

① 상담신청	국민신문고 ▶ 민원상담·안내
② 상담관지정	민원 내용에 따른 상담관 지정
③ 사실관계조사	관계 법령 검토 및 자료조사로 사실관계 확인
④ 결과확인	국민신문고 ▶ 민원상담 신청결과