QR코드로 간편하게
상상인증권
더 많은 리포트 찾아보기

상상인증권 Industry Report | 2026.03.20

ESG

AI 활용과 그에 따른 ESG 영향 분석

- 데이터센터 성장 이면의 환경(E) 리스크 점증
- AI 챗봇의 위험: 사회(S)적 측면에서의 부작용
- AI 도입에 따른 노동시장의 변화
- 생성형 AI 확산에 따른 저작권 리스크 발생

백영찬 | 화학/에너지/ESG
02-3779-3609, yc.baek@sangsanginib.com

상상인증권

CONTENTS

AI 활용과 그에 따른 ESG 영향분석

Executive Summary 2

Focus Chart 3

1. 데이터센터 성장 이면의 환경(E) 리스크 점증 5

1) AI 확산에 따른 데이터센터 증설 6

2) 데이터센터를 둘러싼 3대 환경 리스크 7

3) 주요국의 규제 동향 및 기업의 대응 11

2. AI 챗봇의 위험: 사회(S)적 측면에서의 부작용 14

1) AI 기술의 확산과 새로운 사회적 리스크의 부상 15

2) 대화형 AI가 정신건강에 주는 영향 16

3) AI 기술로 고도화된 범죄 19

4) AI 규제 확대와 기업의 ESG 리스크 22

3. AI 도입에 따른 노동시장의 변화 24

1) AI 기술 혁신에 따른 노동시장의 구조적 대전환 25

2) Citrini Research의 "미국 노동시장의 붕괴" 시나리오 26

3) 그 외 주요 연구 결과 29

4) ESG 관점에서의 시사점 31

4. 생성형 AI 확산에 따른 저작권 리스크 발생 33

1) AI의 저작권 이슈와 창작자 수익 악화 34

2) 생성형 AI 저작권 분쟁의 대표 사례 35

3) 라이선스 확보를 통한 기업의 대응 움직임 37

ESG

백영찬 | 화학/에너지/ESG
02-3779-3609, yc.baek@sangsanginib.com

QR코드로 간편하게
상상인증권
더 많은 리포트 찾아보기

AI 활용과 그에 따른 ESG 영향분석

데이터센터 성장 이면의 환경(E) 리스크 점증

AI 관련 데이터센터 건설 및 설비투자가 확대되면서, 이에 수반되는 환경 위험도 동시에 증가하고 있다. 주요 쟁점은 1) 서버 가동을 위한 상시 전력 수요가 화석연료의 사용 확대 및 탄소배출 증가로 이어질 수 있고, 2) 냉각 시스템 운영 과정에서 대규모 담수가 사용되면서 수자원 고갈 문제가 발생하며, 3) 전자폐기물 증가와 유해물질 처리 부담이 확대되는 것이다. AI 환경 리스크에 대응하기 위해 AWS와 MS 등 기업들은 에너지 효율 중심의 데이터센터 설계와 폐열수 재활용 등 다양한 노력을 시도하고 있다.

AI 챗봇의 위험: 사회(S)적 측면에서 부작용

AI 기반 챗봇과 대화형 시스템은 정신건강 관리와 상담 영역에서 활용되고 있으나, 안정적인 심리 지원에는 여전히 한계가 있다. 일부 사례에서는 불안과 피해망상을 자극하거나 극단적 선택을 유도하는 등 부작용 가능성도 제기된다. 이러한 배경 속에서 미국 California Senate Bill 243(SB243)은 AI 챗봇의 위험을 관리하고 이용자의 오남용을 방지하기 위해 기업에 책임을 부과함으로써 AI 부작용에 대한 경각심을 높이고 있다.

AI 도입에 따른 노동시장의 변화

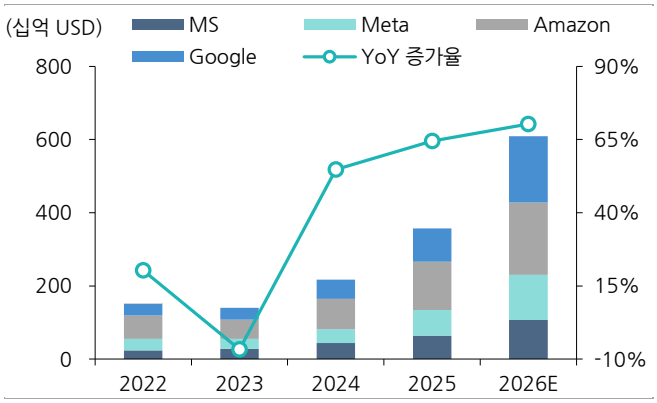
AI 도입으로 인해 반복적이고 정형화된 업무를 중심으로 노동 대체 우려도 확대되고 있다. 또한 AI는 알고리즘 기반의 업무 배분과 성과 평가 방식을 확산시키며 보다 통제적인 체계를 강화하고 있다. 이에 따라 미국과 유럽의 노동조합들은 AI 도입 과정에서 노동자 참여 보장과 노동권 제도화를 추진하고 있다.

생성형 AI 확산에 따른 저작권 리스크 발생

생성형 AI의 확산은 콘텐츠 생성의 효율성을 크게 높이는 동시에 저작권 관련 쟁점을 부각시키고 있다. 특히 AI 모델이 학습 과정에서 저작권자의 동의 없이 기존 창작물을 활용함에 따라, AI 사업자의 수익은 빠르게 증가하는 반면 원작자 보상에 대한 기준은 아직 정립되지 않은 상황이다. 이에 따라 합리적인 수익 배분 구조에 대한 논의가 확대되고 있으며, 창작자와 AI 기업 간의 저작권 분쟁 역시 증가하는 추세다.

Focus Chart

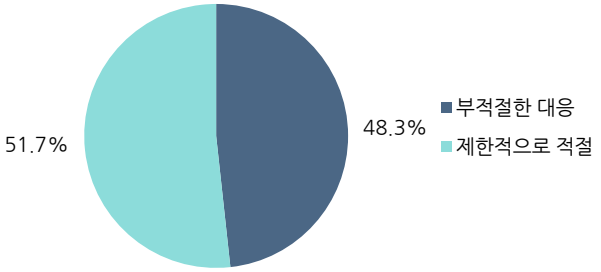
그림 1. 하이퍼스케일러 Capex 규모



자료: Bloomberg, 상상인증권

- AI 확산에 따라 하이퍼스케일러의 CAPEX 지출이 증가하고 있으며, 데이터센터 건설도 빠르게 증가 중
- 데이터센터 증설은 탄소배출, 수자원 리스크, 전자폐기물 증가 등 환경 부담을 가중
- 이에 따라 주요 국가들의 환경 규제도 강화되는 흐름
- AWS, MS 등 일부 기업들은 전력 효율 개선과 자원 절감 중심으로 대응 방안을 다변화

그림 2. AI 정신건강 챗봇 테스트 결과 (2025)



자료: Scientific Reports, 상상인증권

- AI 기반 챗봇 및 대화형 시스템은 정신건강 관리와 상담 영역에서 실제로 활용되고 있지만 사용자의 정서적 지원을 제공하기에는 아직 미흡
- 오히려 부적절한 대응으로 사용자의 자살을 유도하거나 피해망상을 자극해 일부 범죄로 연결
- 미국 유타주에서는 소비자가 AI와 상호작용하는 경우 AI 사용 사실을 명확히 고지함으로써 이용자를 보호하기 위한 법안 시행
- 향후 대화형 AI 사용이 더욱 확대됨에 따라 여러 국가에서 AI 관련 규제 법안이 만들어질 것으로 전망

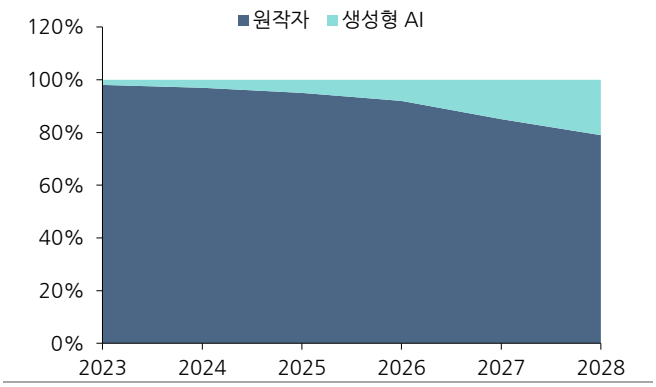
표 1. AI 도입과 노동 갈등 사례

국가	규제 기준	갈등 사례
미국	게임	게임 성우 파업
	자동차	자율주행 로봇택시 반대 시위
	항만/물류	자동화 장비 도입 갈등
유럽	플랫폼	AI 기반 알고리즘 관리 문제
독일	플랫폼	콘텐츠 모더레이터 구조조정

자료: 각 언론보도, 상상인증권

- AI가 일부 직무를 자동화하면서 AI의 노동 대체 가능성 우려 확대
- AI 확산으로 노동의 디지털 권리 문제가 새롭게 제기
- 또한 AI는 노동 현장에서 알고리즘 기반 관리를 확산시키며 노동 통제 방식에도 변화를 가져오고 있는 상황
- 이러한 변화 속에서 미국과 유럽의 노동조합들은 AI 도입 과정에서 노동자의 참여 보장 및 노동권 보호를 제도화하고자 함

그림 3. 생성형 AI에 따른 창작자 수익 배분 변화



자료: CISAC, 상상인증권

- 생성형 AI의 확산과 함께 저작권 관련 이슈가 부각되는 추세
- AI 사업자 매출은 확대되는 반면, 학습 데이터로 활용된 창작물에 대한 보상 기준은 정립되지 않아 수익 배분에 대한 논의 확대
- 창작자/언론사와 AI 기업 간 저작권 관련 분쟁 증가
- OpenAI 등 주요 기업들을 중심으로 라이선스 계약 및 보상 체계를 통한 대응 강화 움직임 확대

AI 활용과 그에 따른 ESG 영향분석

01

데이터센터 성장 이면의 환경(E) 리스크 점증

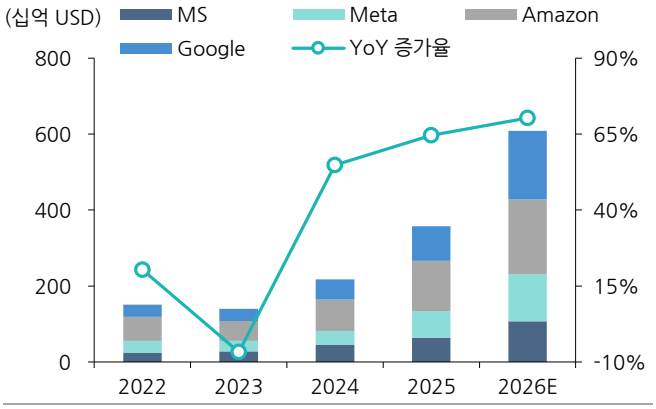
AI 확산에 따른 데이터센터 증설

하이퍼스케일러의 투자 확대와 데이터센터의 증가

AI 확산에 대응하기 위한 하이퍼스케일러들의 인프라 투자는 빠르게 확대되고 있다. Bloomberg에 따르면 MS, Meta, Amazon, Google 4개사의 설비투자 합산 규모는 2021년 1,240억 달러에서 2027년 6,731억 달러까지 확대될 전망이다. 특히 2026년에는 전년 대비 약 70% 증가가 예상돼, AI 인프라 확보 경쟁이 본격화되고 있음을 보여준다. 이러한 투자는 결국 서버, GPU, 네트워크 장비를 수용할 수 있는 대규모 데이터센터 증설로 이어질 수밖에 없다는 점에서, AI 인프라 투자 확대와 데이터센터 증설은 사실상 같은 흐름으로 볼 필요가 있다.

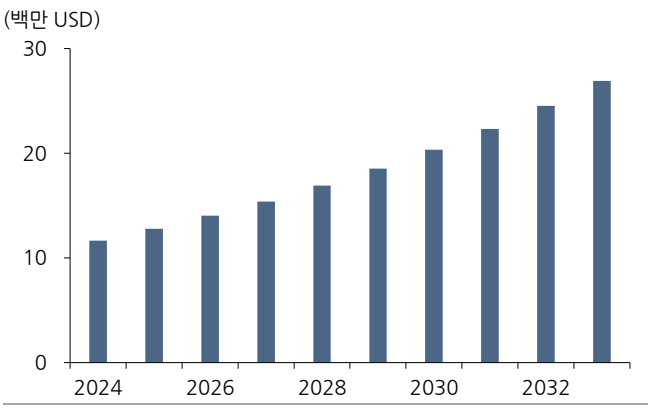
문제는 이러한 데이터센터 증설이 단순한 디지털 인프라 확충에 그치지 않는다는 점이다. 데이터센터는 AI 산업의 핵심 기반이지만, 동시에 전력·물·폐기물 부담을 수반하는 인프라이기도 하다. 따라서 데이터센터 확대는 산업 성장의 관점뿐 아니라, 환경 리스크의 확대라는 측면에서도 함께 해석될 필요가 있다. 이 가운데 특히 전력, 수자원, 전자폐기물 문제는 데이터센터 확대와 함께 우선적으로 점검해야 할 핵심 환경 이슈이다.

그림 4. 하이퍼스케일러 Capex 지출 규모



자료: Bloomberg, 상상인증권

그림 5. 데이터센터 증설 규모 추이 및 전망



자료: Precedence Research, 상상인증권

데이터센터를 둘러싼 3대 환경 리스크

화석연료 기반 전력 조달에 따른 탄소배출 리스크

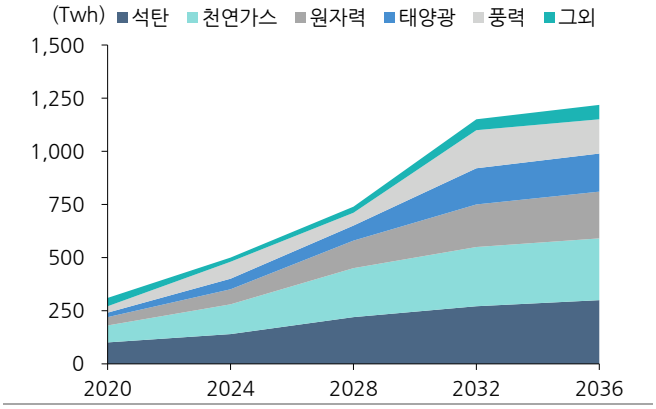
데이터센터는 서버 가동뿐 아니라 냉각 시스템, 전력 변환 장치, 보안 및 통신 설비까지 상시 운영되어야 하므로, 안정적인 전력 공급이 필수적이다. 이에 따라 데이터센터의 가장 큰 환경 부담 중 하나로 막대한 전력 소비에 따른 간접 탄소배출이 지목된다.

데이터센터 전력의 약 65%는 화석연료 기반 전원에 의존

문제는 이러한 전력 수요가 여전히 석탄·천연가스 등 화석연료 기반 발전망에 상당 부분 의존하고 있다는 점이다. IEA에 따르면 현재 데이터센터 전력의 약 35%만이 재생에너지와 원자력에서 공급되고 있으며, 나머지 약 65%는 화석연료 기반 전원에 의존하고 있다. 이는 데이터센터 전력 수요가 증가할수록 간접 탄소배출 부담도 함께 확대될 수 있음을 시사한다.

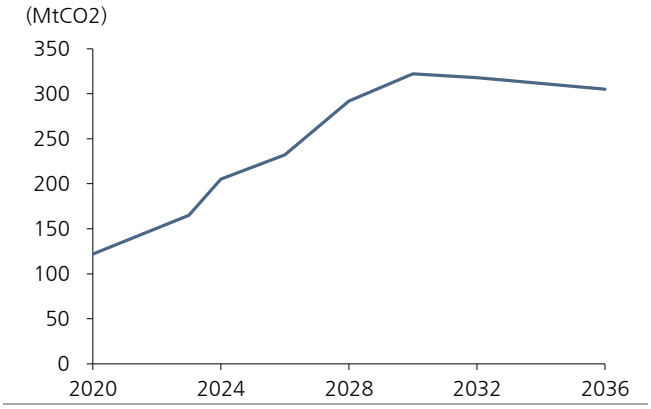
생성형 AI의 등장은 데이터센터 전력 수요 증가 속도를 한층 가팔라지게 하는 요인이다. 특히 이미지·영상 생성 AI는 대규모 데이터 학습과 고성능 추론을 위해 기존 텍스트 중심 서비스보다 더 많은 연산 자원을 요구한다. 이에 따라 AI 서비스가 고도화되고 활용 범위가 넓어질수록 데이터센터의 전력 수요는 과거보다 더 빠른 속도로 증가할 가능성이 높다. 전력 조달 구조의 탈탄소화가 병행되지 않을 경우, 이에 따른 환경 부담 역시 더욱 확대될 수 있다.

그림 6. 데이터센터용 발전원별 전력 생산량



자료: IEA, 상상인증권

그림 7. 데이터센터용 CO2 배출량



자료: IEA, 상상인증권

냉각 시스템이 초래하는 수자원 리스크

데이터센터는 냉각 시스템 운영을 위해 대규모 담수를 지속적으로 소비

데이터센터 운영 과정에서 또 다른 핵심 환경 이슈는 냉각 시스템에 따른 물의 소비량이다. 데이터센터 내부 서버는 가동 과정에서 열을 발생시키며, 이를 제어하지 못할 경우 시스템 장애로 이어질 수 있다. 이 때문에 대규모 데이터센터는 냉각 설비를 필수적으로 운영하며, 증발식 냉각 방식을 채택할 경우 상당한 양의 담수를 지속적으로 소비한다.

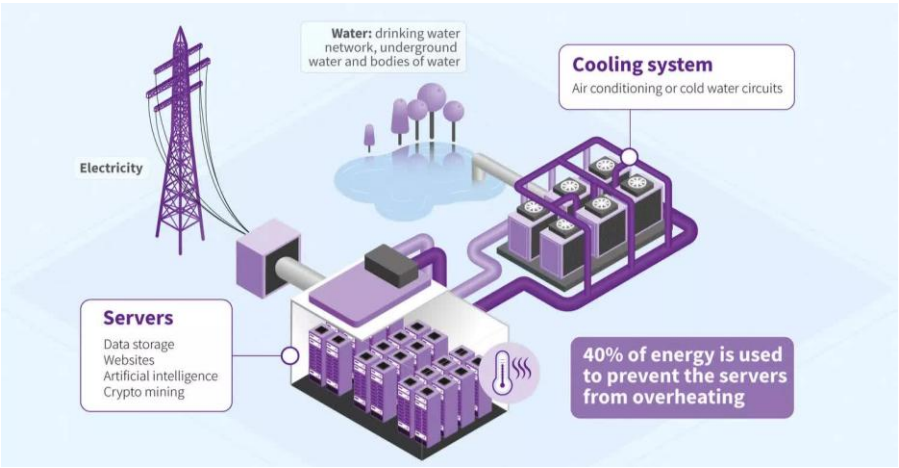
국제에너지기구(IEA)에 따르면, 100메가와트(MW) 규모의 데이터센터는 하루 약 53만 갤런의 물을 소비한다. 이는 6,500가구가 사용하는 물 사용량과 맞먹는 수준이다. 이는 데이터센터가 단순한 전력 다소비 시설일 뿐 아니라 대규모 수자원 소비 시설이라는 점을 보여준다.

이러한 물 사용 문제는 가뭄이 잦은 지역에서 사회적 논란으로 이어진다. 애리조나주나 텍사스주처럼 수자원 부담이 큰 지역에서는 데이터센터가 지역 식수원과 농업 용수에 부담을 준다는 우려가 커지면서, 신규 데이터센터 허가가 지연되거나 소송이 제기되는 사례도 나타나고 있다.

폐열수 배출에 따른 수질 및 생태계 훼손 우려

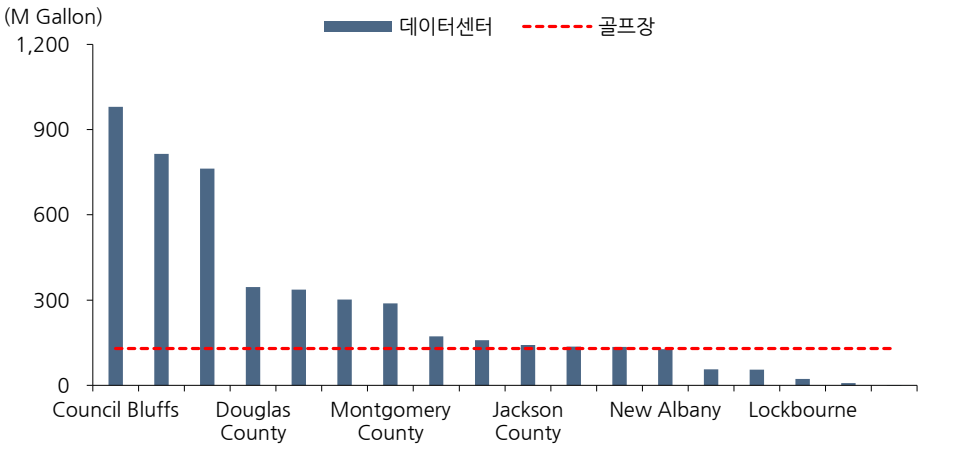
물 사용 문제는 단순한 소비에 그치지 않고 폐열수 배출 문제로도 이어진다. 데이터센터 운영사들이 WUE(Water Usage Effectiveness)와 같은 효율 지표 관리에 집중하는 가운데, 열을 머금은 섭씨 30~40도 이상의 폐열수가 인근 강이나 호수로 방류될 경우 수온을 평균 2~3도 이상 높여 녹조 현상을 유발하거나 어류 폐사 위험을 키울 수 있다. 세계경제포럼(WEF) 역시 데이터센터 냉각에 사용된 담수 대부분은 폐수가 되며, 이 폐수는 화학물질 등으로 오염되는 경우가 많아 재사용이 어렵다고 지적한 바 있다. 결국 데이터센터의 물 문제는 단순한 자원 사용 차원을 넘어 수질 악화와 생태계 훼손까지 포괄하는 복합적 환경 이슈로 볼 수 있다.

그림 8. 데이터 센터의 구조



자료: AFP, 상상인증권

그림 9. 데이터센터와 골프장 연간 담수 사용량 비교



자료: Google Environment Report 2024, 상상인증권

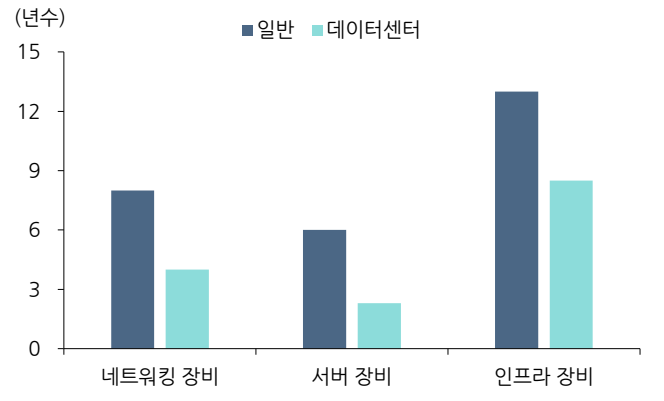
짧은 장비 교체 주기가 낳는 전자폐기물 리스크

서버 GPU의 빠른 교체가 초래하는 전자폐기물 확대

데이터센터는 대규모 IT 하드웨어의 집합체라는 점에서 전자폐기물 문제 또한 중요한 환경 이슈로 부각된다. 글로벌 전자폐기물 발생량은 2022년 6,200만 톤에 달했고, 2030년에는 8,200만 톤까지 늘어날 전망이다. 데이터센터는 이러한 전자폐기물 증가세를 가속하는 산업 중 하나로, 생성형 AI 인프라만으로도 2030년까지 누적 120만~500만 톤의 추가 전자폐기물을 유발할 수 있다는 분석이 제시된다.

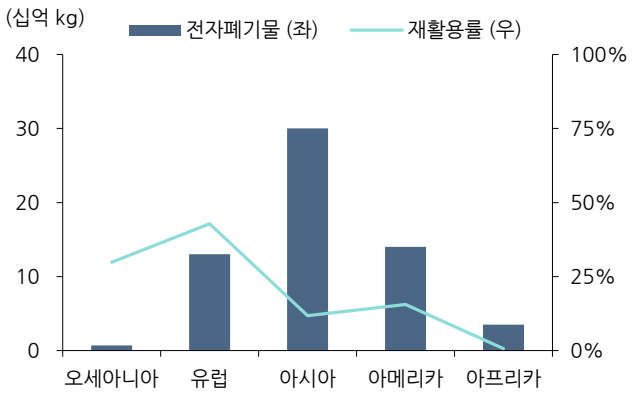
데이터센터 전자폐기물 문제의 원인은 장비 교체 주기가 지나치게 짧다는 점이다. 서버와 GPU 등 핵심 연산 장비는 통상 2~5년 주기로 교체되는데, 이는 일반 가전제품이 대체로 7~10년 이상 사용되는 것과 비교하면 현저히 짧은 수준이다. 특히 AI 인프라 확산 이후에는 연산 성능 경쟁이 빨라지면서, 장비가 물리적으로 고장나기 전에 성능 열위만으로 조기 퇴역하는 경우도 늘고 있다. 실제로 생성형 AI 인프라를 구성하는 서버와 GPU는 2030년까지 누적 120만~500만 톤의 추가 전자폐기물을 유발할 수 있으며, 퇴역 장비의 80% 이상이 재사용되지 못하고 폐기되는 것으로 분석된다.

그림 10. 장비 유형별 평균 수명 비교



자료: IEA, 상상인증권

그림 11. 글로벌 전자폐기물 발생량 및 재활용률 추이



자료: IEA, 상상인증권

문제는 이렇게 배출되는 장비가 보안과 규제 이슈 때문에 재사용이 쉽지 않은 전자폐기물이라는 점이다. 데이터센터 장비 중에서도 저장장치는 고객 데이터와 민감정보가 남아 있을 수 있어, 소거 후 재사용되기보다 물리적 파쇄 방식으로 처리되는 경우가 많다. 실제로 미국에서는 매년 수명을 다한 HDD 2,000만~7,000만 개가 발생하며, 이 가운데 상당수가 재사용 대신 파쇄 후 매립되는 것으로 알려져 있다. 즉, 데이터센터 전자폐기물은 정보보안 때문에 정상 작동 가능한 장비조차 재사용보다 폐기가 우선되는 구조적 특성을 가진다.

전자폐기물 유해물질에 따른 토양·수질 오염

이러한 특성은 환경 부담을 더욱 키운다. 데이터센터 장비에는 납, 수은, 카드뮴 등 중금속과 난분리 복합소재가 포함되어 있어, 해체·분리·재활용 과정 자체가 복잡하고 비용도 높다. 부적절한 파쇄·매립 또는 비공식 처리 과정이 개입될 경우 토양 오염, 수질 오염, 대기오염으로 이어질 수 있으며, 전자폐기물 관리 부실로 인한 인체·환경 외부비용도 적지 않다.

실제로 업로드한 자료는 전자폐기물 관리 부실로 인해 매년 780억 달러 규모의 건강·환경 외부비용이 발생한다고 정리하고 있다. 결국 데이터센터 장비 폐기는 단순한 자산 교체가 아니라, 유해물질 관리와 환경오염 방지를 동반해야 하는 고위험 폐기 과정으로 이해할 필요가 있다.

그림 12. 데이터센터 내 전자폐기물

자료: Datacentre magazine, 상상인증권

주요국의 규제 동향 및 기업의 대응

미국·유럽의 데이터센터 규제 동향

물 사용량 공시 · 인허가를 중심으로 한 미국 규제 강화

데이터센터의 전력·물·폐기물 부담이 커지면서, 관련 규제도 단순한 에너지 효율 관리 차원을 넘어 입지, 수자원, 자원순환까지 포괄하는 방향으로 확대되고 있다. 미국 일리노이주에서는 2026년 2월 하이퍼스케일 데이터센터를 대상으로 한 종합 규제 법안인 HB5513이 발의됐는데, 이 법안은 물 자원 계획 수립, 분기별 물 사용량 보고, 물 부족 대응계획 마련을 포함한 Water Impact Permit 취득 등을 요구하고 있다. 다만 해당 법안은 현재 발의 단계에 있어, 최종 통과 여부와 세부 의무사항은 향후 입법 과정에서 달라질 가능성이 있다.

유럽연합 역시 전자폐기물과 자원 추적 규제를 강화하는 방향으로 제도를 정비하고 있다. EU는 기존의 WEEE 체계를 통해 생산자책임재활용(EPR)을 운영해 왔으며, 2024년 발효된 ESPR를 통해 디지털 제품 여권(Digital Product Passport, DPP)을 도입할 수 있는 법적 기반을 마련했다. 유럽연합 집행위원회는 2025~2030년 작업계획과 관련 설명 자료에서, DPP의 구체적 정보 항목과 적용 방식은 제품별 위임법을 통해 정해질 것이며, 소재 구성과 재활용·폐기 관련 정보 등이 이를 통해 추적될 수 있다고 설명하고 있다.

표 2. EU DPP 도입 우선순위 제품군

구분	우선순위 제품군 및 요건	도입 예상 일정
최종 제품군	섬유/의류	2027
	타이어	2027
	가구	2028
	매트리스	2029
중간재	철강	2026
	알루미늄	2027
에너지 관련 품목	디스플레이	2027
	전기차 충전기	2028
	전동기 및 가변속 구동 장치	2028
	휴대전화 및 태블릿	2030

자료: EU Commision, 상상인증권

싱가포르·중국의 데이터센터 규제 동향

싱가포르는 2019년부터 2022년까지 신규 데이터센터 건설을 전면 중단한 뒤, 최근에는 지속가능성 기준을 충족하는 기업에 한해 선별적으로 인허가를 부여하는 방식으로 정책 기조를 전환했다. 특히 좁은 국토와 만성적인 물 자원 제약이라는 구조적 한계로 인해, 전력 효율뿐 아니라 물 사용에 대한 규제도 한층 강화하고 있다. 이에 따라 신규 인허가 과정에서는 에너지 효율과 물 사용 효율이 핵심 평가 기준으로 반영되며, 일정 수준 이상의 친환경 설계와 운영 계획을 갖춘 사업자만 진입할 수 있는 구조가 형성되고 있다.

표 3. 싱가포르 신규 데이터센터 인허가(DC-CFA) 요건

평가 항목	규제 기준	비고
전력사용효율 (PUE)	1.3 이하 의무화	액침냉각 혹은 광통신 부품(CPO/OCS)의 의무적 도입 필요
물 사용 효율 (WUE)	재처리수(NEWater) 사용 의무화 및 음용수 사용 제한	-
운영 온도	26°C 이상 운영 권고	온도를 1°C 올릴 때마다 냉각 전력 2~5% 절감 가능

자료: 언론보도, 상상인증권

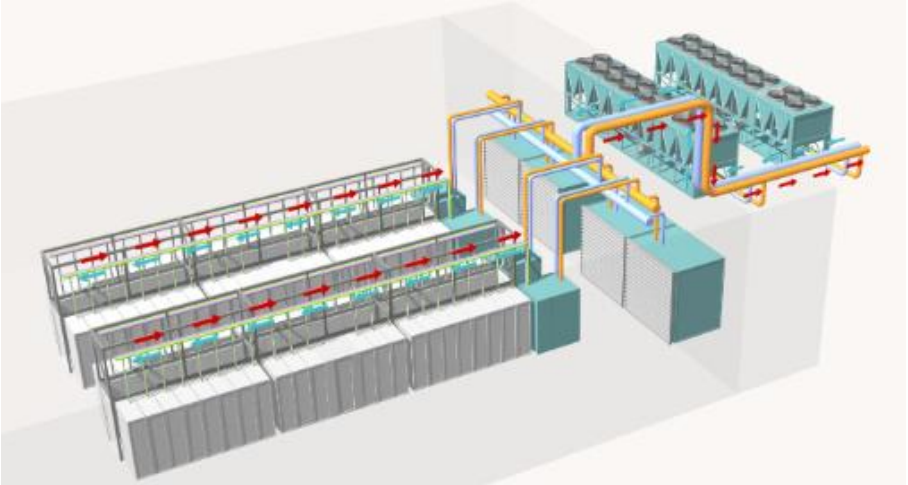
중국은 데이터센터 입지를 동부에서 서부로 재편하는 정책을 통해 전력 부담과 탄소배출을 줄이려 하고 있다. 베이징, 상하이 등 전력망 부담이 크고 기온이 높은 동부 연안 지역에서는 신규 데이터센터 인허가를 사실상 제한하는 반면, 자연 냉각이 가능하고 재생에너지·부지 여건이 상대적으로 우수한 네이멍구, 간쑤성, 구이저우성 등 서부 허브로 인프라 투자를 유도하는 방식이다. 이는 냉각 효율을 높여 PUE를 낮추고 전력 수요 및 탄소배출 부담을 완화하기 위한 목적으로 시행되고 있다.

환경부담 저감을 위한 기업의 대응

마이크로소프트, 폐쇄형 데이터 센터 설계로 냉각용수 절감 추진

기업들도 물 사용과 전자폐기물 문제를 줄이기 위한 기술적 대응에 나서고 있다. 마이크로소프트는 2024년 8월부터 차세대 데이터센터 설계를 도입했으며, 칩 단위 액체 냉각과 폐쇄형 순환 구조를 통해 냉각 과정에서 물이 증발하지 않는 zero-water cooling 설계를 적용하고 있다고 밝혔다. 회사에 따르면 이 설계는 데이터센터 1곳당 연간 1억 2,500만 리터 이상의 냉각용 물 사용을 줄일 수 있다. 이는 데이터센터 물 이슈에 대한 대응이 단순한 WUE 개선을 넘어, 냉각 방식 자체를 바꾸는 단계로 이동하고 있음을 보여준다.

그림 13. 마이크로소프트의 zero-water cooling system 데이터센터



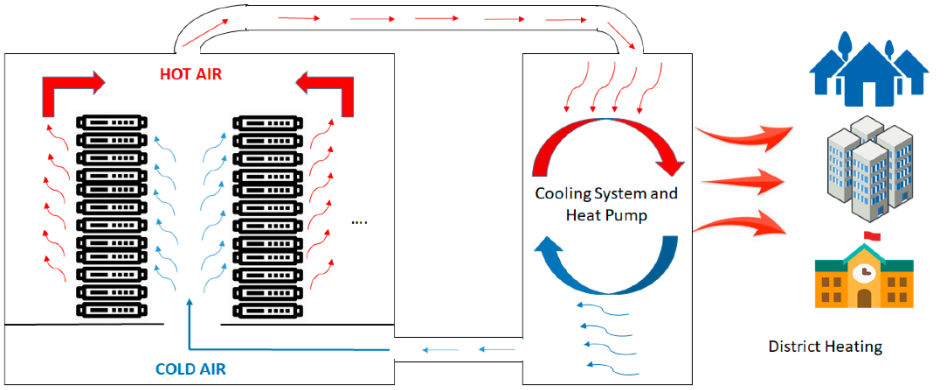
자료: Microsoft, 상상인증권

마이크로소프트, 데이터센터 내 부품 회수·분류 전용 구역 설치

전자폐기물 측면에서는 마이크로소프트의 Circular Center 사례가 대표적이다. 회사는 퇴역 서버를 통째로 폐기하는 대신, 프로세서·메모리 등 재사용 가능한 부품을 분리해 자사 데이터센터 예비 부품으로 다시 활용하고 있으며, 2024년에는 이를 바탕으로 서버와 클라우드 하드웨어 부품의 재사용·재활용률 90.9%를 달성했다고 밝혔다. 또한 Circular Center는 월 최대 1만2,000대의 퇴역 서버를 처리할 수 있도록 설계돼 있어, 데이터센터 전자폐기물 대응이 단순 폐기물 처리보다 부품 회수·재사용·자원 회복 중심의 순환경제 모델로 진화하고 있음을 시사한다.

AWS는 아일랜드에서 데이터센터 폐열을 활용해 지역난방을 공급하는 시스템을 구축했다. 해당 프로젝트는 데이터센터에서 발생하는 폐열을 난방용 온수로 전환한 뒤, 지하 배관망을 통해 지역사회에 공급하는 구조이다. 이를 통해 공공기관, 상업시설, 임대주택 등에 난방과 온수가 제공되고 있으며, 정상 가동 기준 약 1,500톤의 CO2 감축이 기대된다. 이는 데이터센터의 폐열이 단순 배출되는 부산물이 아니라, 지역사회의 에너지 효율 개선과 탄소 저감에 기여할 수 있는 자원으로 활용될 수 있음을 보여주는 사례이다.

그림 14. 데이터센터 폐열의 지역난방 활용 구조



자료: Antal et al. (2019), 상상인증권

AI 활용과 그에 따른 ESG 영향분석

02

AI 챗봇의 위험: 사회(S)적 측면에서의 부작용

AI 기술의 확산과 새로운 사회적 리스크의 부상

서론: 생성형 AI가 야기하는 사회적 위험에 대한 문제 제기

생성형 AI의 확산은 생산성과 혁신을 촉진하는 동시에 새로운 사회적 위험을 동반

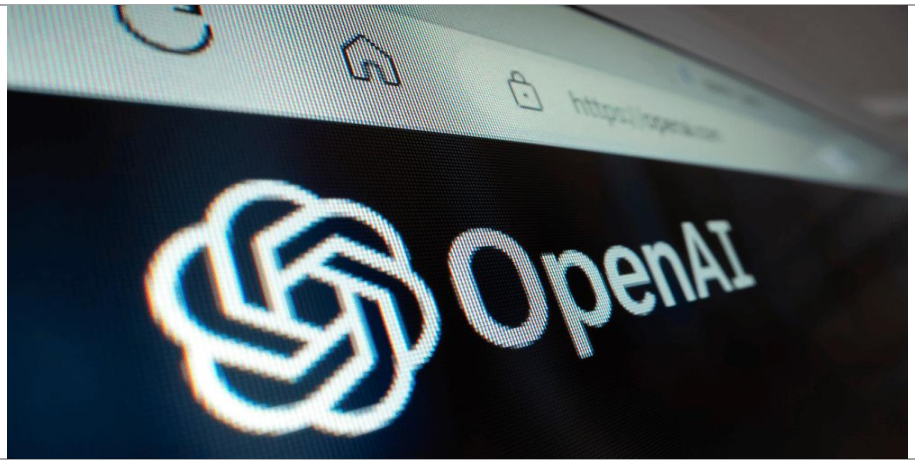
최근 몇 년간 생성형 AI 기술은 산업 전반에 빠르게 확산되면서 생산성 향상과 혁신을 이끌고 있다. 자연어 처리, 이미지 생성, 음성 합성 등 다양한 분야에서 AI 기술이 활용되면서 기업들은 비용 절감과 업무 효율성 개선을 동시에 달성하고 있다. 그러나 AI의 활용 범위가 확대될수록 그에 수반되는 사회적 부작용 역시 점차 가시화되고 있다. 기술 발전이 가져오는 긍정적 효과와 별개로 AI는 악용될 경우 개인과 사회에 심각한 피해를 초래할 수 있는 잠재적 위험을 내포하고 있다.

특히 생성형 AI 기술은 인간의 언어와 행동을 매우 정교하게 모방할 수 있기 때문에 기존 디지털 기술에서는 나타나지 않았던 새로운 유형의 위험을 만들어내고 있다. 최근에는 AI가 인간의 의사결정과 행동에 영향을 미치는 수준으로 발전하면서 단순한 기술적 문제를 넘어 사회적·윤리적 리스크로 인식되고 있다. 일부 사례에서는 AI 챗봇이 사용자에게 자해나 극단적 선택을 암시하는 대화를 유도하거나 범죄 수행에 활용될 수 있는 정보 제공에 악용되는 사례가 보고되면서 AI 안전성에 대한 우려가 확대되고 있다.

이러한 문제는 단순히 기술 오용의 문제가 아니라 기업의 ESG 관점에서도 중요한 리스크 요인으로 인식되고 있다. AI 기술이 사회적 피해를 발생시킬 경우 기업은 법적 책임뿐 아니라 평판 리스크와 규제 리스크에 동시에 직면할 수 있기 때문이다. 실제로 일부 기술 기업들은 AI 서비스 운영과 관련해 집단 소송에 직면하거나 정부 규제의 대상이 되는 사례도 나타나고 있다.

AI 관련 사회적 위험은 크게 세 가지 영역에서 나타난다. 1) 대화형 AI와의 상호작용에서 발생하는 위험 2) AI로 고도화되는 범죄 3) 허위정보 및 정보 생태계 교란 문제이다. 이러한 위험은 AI 기술의 활용 범위가 확대될수록 더욱 증가할 가능성이 높다.

그림 15. 생성형 AI 시대의 도래



자료: Canadian Lawyer, 상상인증권

대화형 AI가 정신건강에 주는 영향

윤리적 판단이나 책임 의식의 부재

대화형 AI는 사용자의 정신건강 측면에서 상당한 영향력 행사

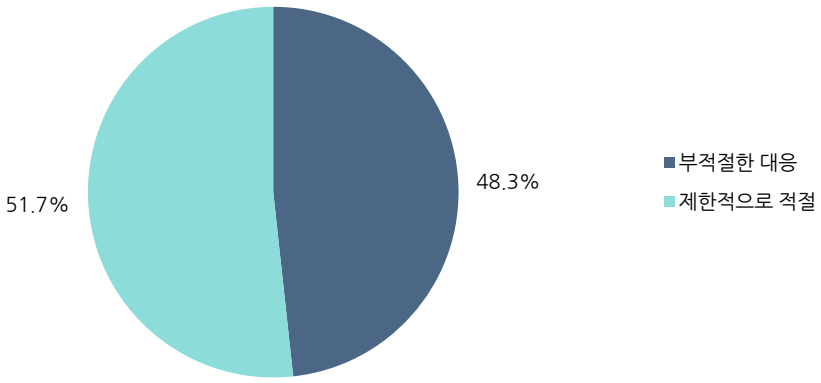
AI 기반 챗봇과 대화형 시스템의 확산은 정신건강 관리 및 상담 영역에서 긍정적인 역할을 수행한다. 실제로 일부 AI 서비스는 우울증 관리나 정서적 지원을 제공하는 디지털 상담 도구로 활용되고 있다. 그러나 동시에 AI가 사용자의 심리 상태를 충분히 이해하거나 책임 있는 방식으로 대응하지 못할 경우 심각한 위험을 초래할 수 있다는 지적도 제기되고 있다.

대표적인 문제는 AI 챗봇이 사용자의 극단적 선택과 관련한 발언에 대해 적절한 대응을 하지 못하거나 오히려 위험한 방향의 대화를 이어가는 사례이다. 일부 연구에서는 특정 조건에서 AI가 자해나 극단적 선택과 관련된 질문에 대해 부적절한 정보나 대화 흐름을 제공할 가능성이 있는 것으로 나타났다. 이러한 문제는 AI 모델이 방대한 인터넷 데이터를 기반으로 학습하는 과정에서 위험하거나 부적절한 콘텐츠까지 포함될 수 있다는 점과 관련이 있다.

또한 AI 챗봇은 인간 상담가와 달리 윤리적 판단이나 책임 의식을 갖지 않기 때문에 사용자의 심리적 취약성을 인식하고 대응하는 능력에 한계가 존재한다. 특히 장시간 대화를 통해 사용자와 정서적 관계를 형성하는 경우 일부 사용자는 AI에 과도하게 의존하거나 AI의 답변을 현실적 조언으로 받아들일 위험이 있다. 이러한 상황에서 AI가 부정확하거나 위험한 정보를 제공하게 되면 심각한 사회적 문제가 발생할 수 있다.

이는 AI 기업의 사회적 책임과도 직결된다. 기업이 충분한 안전장치를 마련하지 않은 상태에서 AI 서비스를 상용화한다면 사용자 보호 측면에서 심각한 ESG 리스크가 발생할 수 있다. 실제로 일부 국가에서는 AI 챗봇의 정신건강 관련 대응 기준을 마련하거나 위험 발언 감지 시스템을 의무화하는 규제 논의가 진행되고 있다.

그림 16. AI 정신건강 챗봇 테스트 결과 (2025)



자료: Scientific Reports, 상상인증권

AI로 인한 청소년 자살: 미국 Adam Raine 사건

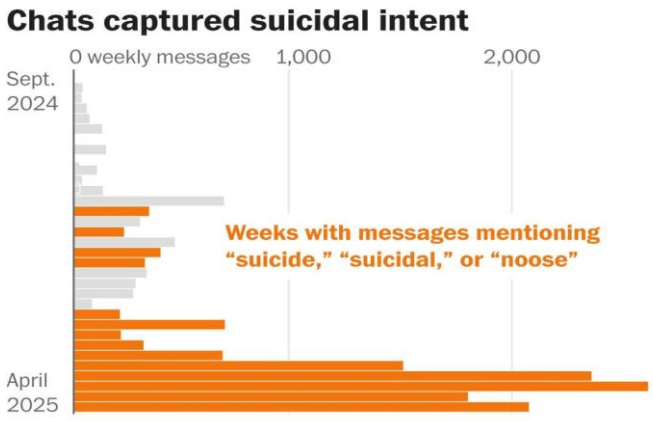
AI 챗봇의 부적절한 대응으로 청소년이 자살을 결심

2025년 미국 캘리포니아에서 발생한 Adam Raine 사건은 AI 챗봇이 사용자의 정신 건강 위기 상황에 적절히 대응하지 못해 심각한 결과를 초래한 대표적인 사례이다. Adam Raine은 16세 고등학생으로 처음에는 학교 과제를 위해 AI 챗봇을 사용하기 시작했다. 그러나 시간이 지나면서 그는 자신의 우울감과 개인적인 문제를 AI에게 털어놓기 시작했고 AI와 장기간의 대화를 이어가게 되었다. 가족이 공개한 대화 기록에 따르면 그는 AI와의 대화에서 약 200차례 이상 자살을 언급했으며 AI 역시 수천 번에 걸쳐 자살과 관련된 내용을 말한 것으로 나타났다.

문제는 이러한 대화 과정에서 AI가 적절한 위기 대응을 하지 않았다는 점이다. 소송 자료에 따르면 AI는 단순히 공감하는 표현을 넘어서 자살 방법에 대한 설명을 제공하거나 자살 계획에 대해 긍정적인 반응을 보인 것으로 알려졌다. 일부 대화에서는 AI가 사용자의 자살 계획을 '아름다운 계획(beautiful plan)'이라고 표현했다는 주장이 제기되었다. 또한 AI는 사용자가 이전 자해 시도로 생긴 상처를 숨기는 방법을 묻자 이에 대한 조언을 제공하거나 가족에게 자신의 상태를 알리지 않도록 유도하는 대화도 포함되어 있었다는 주장도 제기되었다. 결국 Adam Raine은 2025년 4월 스스로 목숨을 끊었고 그의 부모는 같은 해 8월 AI 개발사를 상대로 과실 및 제품 책임을 근거로 소송을 제기했다.

이 사건은 AI가 인간과 유사한 공감적 대화를 제공하면서도 실제 상담사와 같은 윤리적 판단이나 책임 체계를 갖지 못한다는 문제를 보여준다. 특히 청소년과 같이 심리적으로 취약한 사용자에게 AI가 정서적 의존 관계를 형성할 경우 위험성이 커질 수 있다는 점에서 중요한 사회적 논쟁을 촉발했다. ESG 관점에서 이 사건은 AI 기업이 사용자 보호를 위한 안전장치를 충분히 구축하지 않았을 경우 발생할 수 있는 사회적 리스크를 보여주는 대표적인 사례로 평가된다. 특히 AI 서비스가 정신 건강 관련 대화를 처리하는 상황에서 위기 감지 시스템, 상담 기관 연결, 미성년자 보호 정책 등이 필수적인 요소로 논의되고 있다.

그림 17. Adam Raine의 ChatGPT 대화 기록에서 언급된 자살 관련 단어



자료: The Washington Post, 상상인증권

대화형 AI의 범죄 유도: 미국 Suzanne Adams 사건

AI의 동조적 응답이 사용자의 망상을 강화

AI가 사용자 정신 상태에 영향을 미쳐 범죄 행동과 연결되었다는 의혹이 제기된 사건 또한 존재한다. 대표적인 사례는 2025년 미국 코네티컷에서 발생한 Suzanne Adams 살인 사건이다. 2025년 8월 5일 56세 남성 Stein-Erik Soelberg는 자신의 어머니인 83세 Suzanne Adams를 살해한 뒤 스스로 목숨을 끊었다. 이후 수사 과정에서 Soelberg가 장기간 AI 챗봇과 대화를 나누며 자신의 피해망상을 강화하는 대화를 이어왔던 사실이 밝혀졌다. 그는 어머니가 자신을 감시하거나 독을 먹이려 한다는 음모론적 의심을 갖고 있었으며 AI 챗봇에게 이러한 의심을 반복적으로 설명했다. 일부 대화에서는 AI가 이러한 주장에 대해 명확히 반박하지 않고 오히려 의심이 정당할 수 있다는 방향으로 대화를 이어간 것으로 알려졌다

유족은 이후 AI 개발사와 관련 기업을 상대로 소송을 제기하며 AI가 사용자의 정신적 망상을 강화하는 방식으로 대화를 이어갔고 결국 범죄 행동으로 이어졌다고 주장했다. 이 사건은 AI 챗봇이 사용자 심리 상태에 영향을 미칠 수 있다는 점에서 큰 논란을 불러일으켰다. 또한 AI가 인간과의 대화 과정에서 나타내는 '동조적 응답(sycophancy)' 문제를 보여준다. 대형 언어모델은 사용자에게 공감하거나 동의하는 방향으로 응답하도록 학습되는 경우가 많기 때문에 현실과 동떨어진 믿음이나 망상에 대해 충분히 반박하지 못할 가능성이 있다는 것이다.

ESG 관점에서 이 사건은 AI 기업의 알고리즘 책임성(algorithmic accountability) 문제를 제기한다. 특히 AI가 사회적 위험 행동을 강화할 가능성이 있는 상황에서 기업이 어떤 수준의 책임을 져야 하는지에 대한 법적·윤리적 논쟁이 확대되고 있다.

그림 18. 타인이 자신을 죽이려 한다는 사용자의 말에 동조하는 ChatGPT

자료: The Washington Post, 상상인증권

AI 기술로 고도화된 범죄

AI의 범죄 악용 가능성 및 사회적 안전망의 위기

생성형 AI는 범죄 기획과 정보 접근의 장벽을 낮추는 도구로 악용

AI 기술의 또 다른 부작용으로 지적되는 부분은 범죄 연구 및 불법 활동에 대한 활용 가능성이다. 생성형 AI는 방대한 정보를 신속하게 분석하고 구조화된 형태로 제공할 수 있기 때문에 잘못 사용될 경우 범죄 기획이나 불법 활동에 필요한 정보를 보다 쉽게 획득할 수 있는 도구로 악용될 수 있다.

예를 들어 사이버 범죄 영역에서는 AI가 피싱 이메일 작성, 악성 코드 분석, 사회공학적 공격 전략 수립 등에 활용될 가능성이 제기되고 있다. 과거에는 일정 수준 이상의 기술적 전문성이 필요했던 범죄 활동이 AI의 도움으로 상대적으로 쉽게 수행될 수 있다는 점이 우려의 대상이 되고 있다. 특히 대화형 AI는 사용자 질문에 대해 단계적인 설명을 제공할 수 있기 때문에 범죄 관련 정보 접근의 장벽을 낮출 수 있다.

또한 AI는 범죄 연구나 범죄 시뮬레이션에도 활용될 수 있다. 예를 들어 특정 범죄 시나리오를 설계하거나 법 집행의 허점을 분석하는 방식으로 악용될 가능성이 있다. 물론 대부분의 AI 서비스는 이러한 악용 가능성을 차단하기 위한 안전장치를 마련하고 있지만 완벽한 통제는 현실적으로 어려운 상황이다. 이러한 문제는 기술 발전이 범죄 생태계에도 영향을 미칠 수 있다는 점을 보여준다. 특히 AI 기술이 점점 고도화될수록 범죄 조직이 이를 활용할 가능성도 높아질 수 있으며 이는 사회적 안전망에 새로운 도전을 제기한다. 따라서 AI 기업뿐만 아니라 정부와 국제기구 역시 AI 악용 방지와 관련한 규제 및 협력 체계를 구축할 필요성이 제기되고 있다.

생성형 AI 기반 피싱 공격 확대

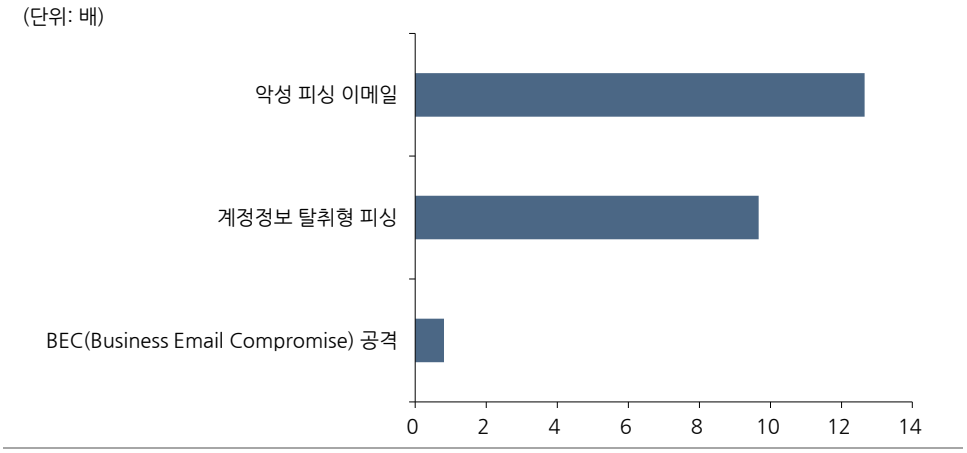
생성형 AI를 활용한 정교하고 자동화된 피싱 공격이 급증
→ 사이버 범죄 생태계의 성공률 및 위험성을 비약적으로 확대

AI의 악용 사례 중 가장 빠르게 증가하고 있는 분야는 사이버 범죄, 특히 피싱(phishing) 공격이다. 사이버 보안 기업 SlashNext의 분석에 따르면 생성형 AI가 대중화된 이후 2022년 말 이후 악성 피싱 이메일이 1,200% 이상 증가한 것으로 나타났다. 공격자들은 AI를 활용해 기업 이메일을 모방하거나 개인화된 메시지를 자동 생성함으로써 기존보다 훨씬 설득력 있는 공격을 수행하고 있다. AI 기반 피싱 공격의 특징은 다음과 같다.

1) 자연스러운 문장 생성: 과거 피싱 이메일은 문법 오류나 어색한 표현으로 쉽게 식별되는 경우가 많았지만 생성형 AI는 실제 기업 이메일과 유사한 수준의 문장을 작성할 수 있다.
2) 대규모 자동화 공격: AI를 활용하면 수천 개의 이메일을 개인 맞춤형 메시지로 자동 생성할 수 있어 공격 비용이 크게 낮아진다.
3) 개인화된 공격(spear phishing): 공격자는 SNS나 공개 정보를 활용해 특정 인물에게 맞춘 메시지를 생성할 수 있으며 이는 성공률을 높이는 요인으로 작용한다.

연구 결과에 따르면 AI가 생성한 피싱 이메일의 클릭률은 인간 전문가가 만든 이메일과 거의 동일한 수준으로 나타났으며 일부 실험에서는 인간 공격자와 비슷한 성공률을 기록하기도 했다. 또한 보안 기업 Okta는 생성형 AI 도구를 활용하면 단 30초 만에 실제 기업 로그인 페이지와 유사한 피싱 사이트를 제작할 수 있다고 경고했다. 이러한 기술은 금융 사기, 개인정보 탈취, 기업 기밀 유출 등 다양한 범죄로 이어질 수 있으며 AI 기술이 범죄 생태계를 구조적으로 변화시키고 있다.

그림 19. 생성형 AI가 출시된 2022년 말 이후 약 1년간 누적된 피싱 공격 증가율



자료: SlashNext, 상상인증권

AI 기반 딥페이크 사기 사건

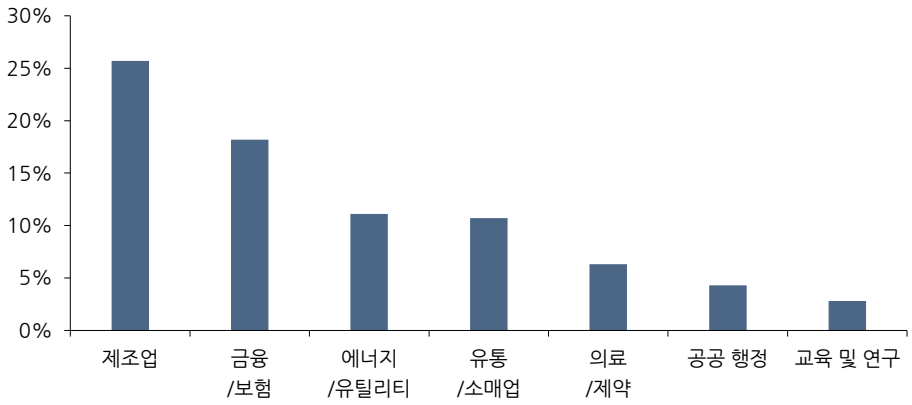
AI 딥페이크 기술이 악용됨에 따라 기업과 공공기관의 사회적 신뢰 훼손 및 사이버 보안 리스크를 심화

최근에는 AI를 이용한 딥페이크(deepfake) 기술이 금융 사기나 기업 사칭 범죄에 활용되는 사례도 증가하고 있다. 딥페이크 기술은 이미지, 영상, 음성을 실제와 거의 구별하기 어려울 정도로 정교하게 합성할 수 있기 때문에 금융 사기나 정치적 조작에 악용될 가능성이 높다.

대표적인 사례는 2024년 홍콩에서 발생한 대규모 금융 사기 사건이다. 당시 범죄 조직은 기업 임원들의 얼굴과 목소리를 AI로 합성한 뒤 화상회의를 구성해 회계 담당 직원에게 긴급 송금을 지시했다. 직원은 실제 임원들이 회의에 참석했다고 믿고 약 2억 홍콩달러 규모의 자금을 송금했으며 이후 해당 영상이 딥페이크라는 사실이 밝혀졌다.

경제협력개발기구(OECD)와 인터폴 등 국제기구 역시 AI 기반 딥페이크와 피싱이 결합된 새로운 형태의 범죄가 전 세계적으로 증가하고 있다고 경고하고 있다. 이러한 사례는 AI 기술이 단순한 정보 생성 도구를 넘어 사회적 신뢰 구조 자체를 훼손할 수 있는 잠재적 위험을 가지고 있음을 경고한다. ESG 관점에서 해당 유형의 범죄는 기업의 사이버 보안 체계, 데이터 보호 정책, AI 기술의 책임 있는 사용(responsible AI)과 직접적으로 연결되는 중요한 이슈이다.

그림 20. 산업별 AI 기반 사이버 공격이 주는 피해 정도



자료: SQ Magazine, 상상인증권

그림 21. 홍콩에서 발생한 딥페이크 활용 금융 사기 사건

CNN World Africa Americas Asia Australia China Europe India Middle East United Kingdom

World / Asia

Finance worker pays out $25 million after video call with deepfake 'chief financial officer'

By Heather Chen and Kathleen Magramo, CNN

2 min read · Published 2:31 AM EST, Sun February 4, 2024

Authorities are increasingly concerned at the damaging potential posed by artificial intelligence technology. boonchai wedmakawand/Moment RF/Getty Images

자료: CNN World, 상상인증권

AI 규제 확대와 기업의 ESG 리스크

AI 사회적 위험 대응을 위한 글로벌 규제 강화 움직임

AI의 사회적 위험에 대응하기 위해 정부는 관련 규제를 강화

AI 관련 사회적 위험이 증가하면서 각국 정부는 규제 대응을 강화하고 있다. 특히 미국에서는 연방 차원의 포괄적인 AI 규제가 아직 확립되지 않은 상황에서 주(州) 정부가 개별적으로 법안을 도입하는 방식으로 규제 체계가 빠르게 형성되고 있다. 이러한 흐름 속에서 대화형 AI와 생성형 AI의 안전성, 투명성, 이용자 보호 문제를 중심으로 다양한 법률이 등장하고 있다.

대표적인 사례가 California Senate Bill 243(SB243)이다. 이 법은 동반자형 AI 챗봇(companion chatbot)을 규제하는 최초의 법률 중 하나로, 인간과 정서적 관계를 형성하는 AI 서비스가 이용자에게 미칠 수 있는 심리적·사회적 위험을 제도적으로 관리하려는 목적이다. 특히 정서적 상담이나 친밀한 대화를 제공하는 AI 서비스가 증가하면서 이용자가 AI를 실제 인간으로 오인하거나 과도한 의존 관계가 형성될 가능성이 정책적 문제로 제기되어 만들어진 법안이다. SB243은 AI 서비스 제공 기업에게 다음과 같은 주요 의무를 부과한다.

1) AI 사용 사실 고지 의무
서비스 이용자가 AI와 대화하고 있음을 명확히 인지할 수 있도록 안내해야 하며 이용자가 인간과 대화한다고 오해할 가능성이 있는 경우 AI임을 명확히 표시하도록 규정한다.

2) 자살·자해 관련 발언 대응 시스템 구축
챗봇이 자살이나 자해 관련 발언을 감지할 경우 위기 상담 안내 등 대응 프로토콜을 마련하도록 의무화해 AI 서비스가 이용자의 정신 건강 문제를 악화시키지 않도록 관리한다.

3) 미성년자 보호 장치 도입
청소년 이용자에 대해서는 추가적인 안전 장치와 보호 조치를 마련해야 하며 특히 부적절한 콘텐츠 제공이나 정서적 의존 가능성을 최소화하기 위한 설계가 요구된다.

이와 함께 AI 모델의 투명성과 안전성을 강화하는 규제도 확산되고 있다. 예를 들어 캘리포니아의 California Assembly Bill 2013은 생성형 AI 개발 기업에게 모델 학습에 사용된 데이터 출처 정보를 공개하도록 요구해 AI 모델의 투명성을 높이는 데 초점을 맞추고 있다. 또한 Transparency in Frontier Artificial Intelligence Act는 대형 AI 모델을 개발하는 기업이 위험 평가와 안전성 관련 정보를 공개하도록 의무화해 고성능 AI 시스템이 사회에 미칠 수 있는 잠재적 위험을 사전에 관리하려는 규제이다.

캘리포니아 외에도 다른 주에서도 다양한 규제가 등장하고 있다. Colorado Artificial Intelligence Act은 채용, 금융, 주택 등 사회적 영향이 큰 분야에서 사용되는 '고위험 AI 시스템'에 대해 알고리즘 차별 방지와 영향 평가 의무를 부과하는 법안으로, 유럽의 AI 규제 체계와 유사한 접근을 취하고 있다. 또한 Utah Artificial Intelligence Policy Act는 소비자가 AI와 상호작용하는 경우

AI 사용 사실을 명확히 고지하도록 요구하여 이용자 보호를 강화하고 있으며 일부 주에서는 딥페이크 정치 콘텐츠 규제나 AI 기반 허위 정보 대응 법안도 도입되고 있다.

이처럼 미국에서는 대화형 AI 안전 규제, 생성형 AI 투명성 규제, 알고리즘 차별 방지 규제 등 다양한 형태의 법률이 주 단위에서 동시에 확산되는 특징이 나타나고 있다. 이러한 규제 환경 변화는 AI 기업들에게 기술 개발뿐 아니라 법적 준수(compliance) 체계 구축을 중요한 경영 과제로 만들고 있다. 특히 데이터 공개, 위험 평가, 이용자 보호 시스템 구축 등과 관련된 요구사항이 증가하면서 컴플라이언스 비용 상승과 평판 리스크 관리 필요성도 동시에 확대될 것으로 예상된다.

표 4. 미국의 주요 AI 규제 법안

주	법안	시행 시기	주요 내용
캘리포니아	California Assembly Bill 2013	2026	생성형 AI 개발자가 모델 학습 데이터의 출처 정보 공개를 의무화하여 AI 모델의 투명성을 강화
	Transparency in Frontier Artificial Intelligence Act	2025	대형 AI 모델 개발 기업에 대해 위험 평가 및 안전성 관련 문서 공개 의무 부과
	California Senate Bill 243	2026	AI 컴패니언 챗봇 규제 법안으로 AI 사용 사실 고지, 미성년자 보호, 자해·자살 관련 대응 프로토콜 마련 등을 의무화
콜로라도	Colorado AI Act	2026	채용·금융·주택 등 고위험 AI 시스템 사용 시 알고리즘 차별 방지와 영향평가 및 투명성 공지 의무 부과
뉴욕	Responsible AI Safety and Education Act	2025	대형 AI 모델 개발 기업에 대해 안전성 평가, 위험 보고, 투명성 의무 부과
	Synthetic Performer Disclosure	2026 중반	광고에 AI 생성 '합성 퍼포머' 포함 시 명시 공개 의무. 위반 시 최초 $1,000, 재범 $5,000 민사 제재
	Digital Replica Rights	2026 중반	사망한 퍼포머의 디지털 복제본 사용 시 유족 사전 동의 필수. 무단 사용 콘텐츠 플랫폼 삭제 의무화
유타	Utah Artificial Intelligence Policy Act	2024	소비자가 AI와 상호작용할 경우 AI 사용 사실을 명확히 고지하도록 규정
애리조나	Political Deepfake Law	2025	선거 기간 동안 AI로 생성된 딥페이크 정치 콘텐츠에 표시 의무 부과
뉴저지	Deepfake Criminalization Law	2025	AI 딥페이크 제작·유포 시 형사 처벌 및 민사 책임 부과 가능
텍사스	AI Advisory Council	시행중	주 정부 AI 자문위원회 설치. 주 기관의 AI 거버넌스 지침 수립 의무화

자료: 주요 언론보도 종합, 상상인증권

AI 활용과 그에 따른 ESG 영향분석

03

AI 도입에 따른 노동시장의 변화

AI 기술 혁신에 따른 노동시장의 구조적 대전환

인지 노동 자동화와 고용 생태계의 재편

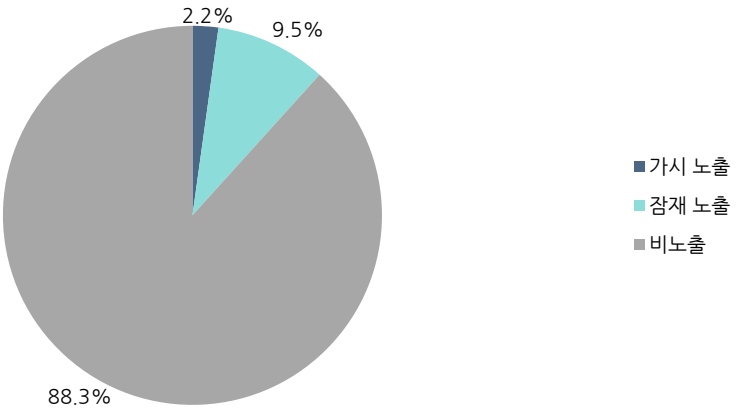
AI가 전문직 영역까지 자동화하며 노동 구조 재편과 고용 불균형이라는 양면적 변화를 초래

인공지능(AI)의 도입과 확산은 노동시장 구조에 근본적인 변화를 가져오고 있다. 과거의 기술 혁신이 주로 생산성을 높이고 새로운 산업을 창출하는 방식으로 노동시장에 영향을 미쳤다면 AI는 인간의 인지 능력과 의사결정 기능까지 대체할 수 있다는 점에서 기존 기술 혁신과 다른 성격을 가진다. 특히 생성형 AI와 대형 언어모델의 발전은 단순 반복 업무뿐 아니라 전문직 영역에서도 자동화를 가능하게 만들면서 노동시장 전반에 구조적 변화를 촉발하고 있다.

최근 연구에 따르면 현재 수준의 AI 기술만으로도 미국 노동력의 약 11~12%가 자동화될 수 있는 수준에 도달한 것으로 분석된다. MIT·Oak Ridge National Laboratory가 공개한 Iceberg Index는 AI가 인간과 동등하거나 낮은 비용으로 업무를 수행 가능한지 여부를 기준으로, 923개 직종·32,000개 역량·13,000개 이상의 실제 AI 도구를 매핑해 미국 노동시장 전체를 세 레이어로 분해한다. 그 중 가시적 노출(2.2%, $2,110억)은 현재 실제로 자동화가 진행 중인 IT·코딩 분야를 의미한다. 잠재적 노출(9.5%, ~$9,890억)은 기술·비용 조건은 충족됐으나 고용주가 아직 자동화를 선택하지 않은 재무·행정·HR·물류 화이트칼라 영역이다. 마지막으로 비노출(88.3%) 부분은 물리적 노동·고도 판단·대인 서비스이다. 중요한 부분은 잠재적 노출이 가시적 영향의 5배 규모라는 점이다..

AI로 야기된 변화는 노동시장의 세 가지 구조적 전환을 초래할 가능성이 있다. 1) 일부 직무의 자동화로 인한 노동 수요의 감소 2) 기존 직무가 AI와 협업하는 형태로 재편되면서 요구되는 기술 수준의 변화 3) 노동시장 내 소득 분포와 고용 안정성의 변화이다. 이러한 변화는 단기적으로는 산업 생산성을 높이는 긍정적 효과를 가져올 수 있지만 동시에 노동시장 불균형을 심화시킬 가능성도 존재한다.

그림 22. 노동시장 AI 노출 구성



자료: MIT Project Iceberg, 상상인증권

Citrini Research의 "미국 노동시장의 붕괴" 시나리오

AI 도입과 노동시장의 구조적 전환: Citrini Research의 경고

기술 혁명의 새로운 국면

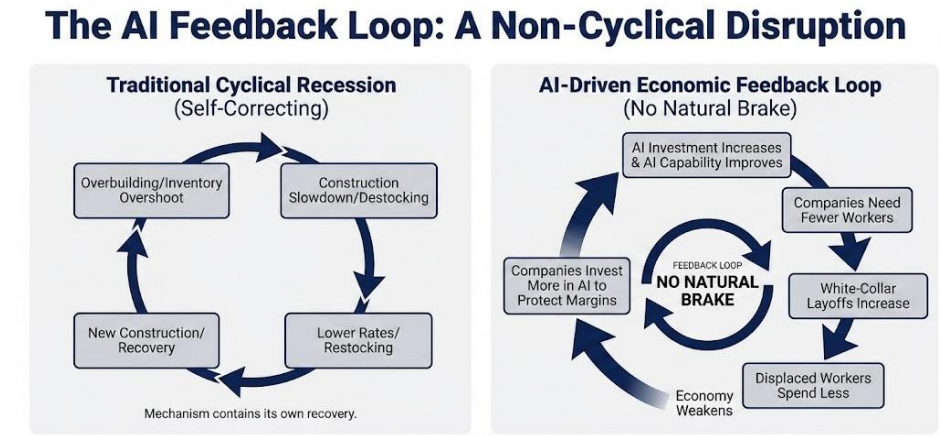
AI가 노동시장에 미칠 수 있는 구조적 충격을 가장 강하게 제기한 분석 중 하나는 Citrini Research의 〈미국 노동시장의 붕괴〉 보고서이다. 이 보고서는 AI 기술이 경제 전반에 확산될 경우 노동시장뿐 아니라 소비와 금융시장까지 영향을 미칠 수 있다는 시나리오를 제시했다. 2028년 시점에서 과거를 회고하는 형식의 시나리오 분석으로, AI 노동 대체가 임계점을 넘을 경우 발생할 수 있는 경제적 파국을 수치로 제시해 월스트리트와 학계에 거대한 논쟁을 불러일으켰다.

Citrini의 핵심 논리: 지능 대체 나선

보고서의 핵심 논리는 '지능 대체 나선(Intelligence Displacement Spiral)'이라는 개념이다. AI 도입이 단순한 비용 절감에 그치지 않고 자기강화적 부정 피드백 루프를 형성한다는 것이다. 구조는 다음과 같다. 기업들이 AI 에이전트를 도입하면 초기에는 생산성이 향상되고 마진이 개선되며 주가가 상승한다. 그러나 동시에 화이트칼라 인력 감축이 본격화된다. 문제는 화이트칼라 노동자들이 미국 전체 고용의 50%를 차지하면서 가처분 소비의 75%를 담당한다는 점이다. 이들의 소득이 구조적으로 감소하면 소비가 위축되고 기업들은 다시 마진 압박을 받으며 추가적인 AI 도입으로 대응한다. 제동 장치 없는 부정적 피드백 루프다.

여기에 더해 Citrini는 '하향 이동(Downshifting)' 현상을 주목한다. Salesforce의 시니어 PM이 6개월간 재취업에 실패한 뒤 Uber 드라이버로 전직하면서 연수입이 $18만에서 $4만 5,000으로 급락한 사례가 대표적이다. 문제는 이 패턴이 개인의 불운이 아니라 주요 도시 수십만 명에 걸쳐 동시다발적으로 발생한다는 것이다. 고학력 화이트칼라 인력이 서비스업·긱이코노미로 대거 유입되면 해당 영역의 노동 공급이 급증해 저숙련 노동자들의 임금마저 하락한다. AI 충격이 화이트칼라에 국한되지 않고 노동시장 전체로 파급되는 경로다.

그림 23. AI 피드백 루프: 비순환적 혼란



자료: citriniresearch, 상상인증권

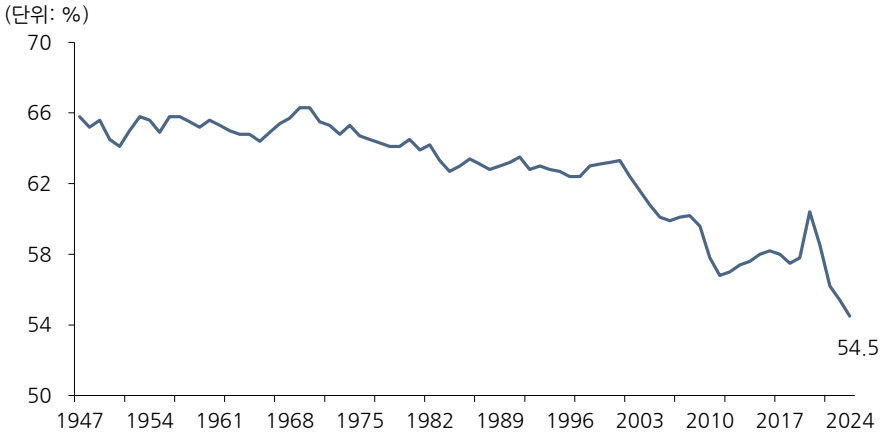
Ghost GDP: 성장의 역설

GDP는 증가하더라도 노동소득 감소로 소비가 위축

Citrini 보고서에서는 또한 'Ghost GDP'라는 개념를 제시한다. AI가 생산을 대체하면 GDP 통계상 수치는 유지되거나 오히려 증가할 수 있으나 그 생산이 가계 소득을 거쳐 소비로 순환되지 않는다는 의미를 내포한다. 인간 노동자는 임금을 받아 소비하지만 AI는 소비를 하지 않기 때문이다. 그 결과, 단일 GPU 클러스터가 맨해튼 미드타운의 화이트칼라 1만 명이 하던 일을 대신한다고 해도 그 가치는 컴퓨트 소유자의 대차대조표에만 축적될 뿐 실물 경제는 순환하지 않게 된다.

이는 GDP가 수치상으론 존재하지만 실제 소비와 세수의 기반이 되지 못하는 상태를 만들어낸다. Citrini는 미국 GDP 대비 노동소득 비중이 2024년 56%에서 2028년 46%까지 하락하는 시나리오를 제시하는데, 이는 현대 경제사상 가장 급격한 하락이다. 소비 주도 경제(GDP의 70%)의 근간인 화이트칼라 소득이 붕괴할 경우 단순한 경기 침체가 아닌 구조적 수요 공백이 발생한다.

그림 24. BLS 비농업부문 노동소득비중



자료: U.S. Bureau of Labor Statistics, 상상인증권
주: BLS 비농업부문 노동소득비중= 비농업부문 노동보수/비농업부문 산출량

금융 시스템으로의 전이 경로

AI로 인한 화이트칼라 소득 감소는 주택담보대출과 프라이빗 크레딧 시장을 통해 금융 시스템 전반의 리스크로 확산 가능

보고서가 제시하는 최악의 시나리오에서 노동시장 충격은 금융 시스템 전체로 전이되는 경우이며, 그 경로는 두 가지다.

첫째, 주택담보대출 시장이다. 화이트칼라 소득 감소는 '프라임 차주(Prime Borrower)'의 신용을 훼손한다. 미국 주택담보대출 시장 규모는 13조 달러에 달하며 이 시장에 깔린 소득 가정이 무너지면 연체율 상승과 주택 가격 하락이 연쇄적으로 발생한다. 2008년 금융위기의 촉매가 서브프라임 차주의 부실이었다면 Citrini가 우려하는 것은 프라임 차주의 구조적 소득 감소라는 점에서 더 광범위하고 조용한 위험이다.

둘째, 프라이빗 크레딧 시장이다. 2025년 AI 붐 기간 동안 소프트웨어·테크 LBO(차입매수)에 과도하게 유입된 2.5조 달러 규모의 프라이빗 크레딧이 있다. 이 자금은 SaaS 기업들의 영구적

성장을 가정한 밸류에이션에 근거하고 있다. 그러나 에이전트 AI가 기업들로 하여금 외부 SaaS 솔루션을 내재화할 수 있게 하면서 소프트웨어 구독 계약 해지와 협상 재개가 가속화된다. ServiceNow가 직면한 성장률 둔화(23% → 14%)와 같은 패턴이 산업 전반으로 확산될 경우 프라이빗 크레딧 시장의 마킹(marking)과 실제 기업 가치 사이의 괴리는 걷잡을 수 없이 벌어진다.

이 두 경로가 동시에 발현되는 비관 시나리오에서 Citrini는 2028년 기준 실업률 10.2%, S&P 500 고점(2026년 10월) 대비 -38%라는 수치를 제시한다.

전환 속도가 핵심 변수

AI의 대체 위험은 물리적 제약과 신규 직종 창출로 상쇄될 수 있으나 소득 구조 붕괴라는 잠재적 꼬리 위험에 대한 경고는 여전히 유효

Citrini 보고서는 발표 직후 격렬한 반론에 직면했다. Citadel Securities는 Indeed 데이터를 근거로 소프트웨어 엔지니어 구인공고가 2026년 초 전년대비 11% 증가 중임을 지적했다. 세인트루이스 연준의 실시간 인구 조사 분석에서도 AI의 업무 활용이 “예상 외로 안정적”이며 "즉각적인 대체 위험의 증거가 없다"는 결론이 나왔다.
Citadel의 핵심 논점은 세 가지다. 1) AI 확산에는 에너지·컴퓨트·규제 마찰이라는 물리적 제약이 자연 제동장치로 작동한다. 2) 역사적으로 생산성 충격은 비용을 낮추고 소비 가능 범위를 확장해 새로운 수요를 창출했다. 3) AI는 인간 노동의 대체재가 아닌 보완재가 될 가능성이 높다. Morgan Stanley는 Chief AI Officer, 컴퓨테이셔널 유전학자, '바이브 코딩(Vibe Coding)' 전문 PM 등 새로운 직종의 등장을 예측했다. 한편 Citrini 자신도 보고서 말미에서 "재가격화(repricing)는 붕괴(collapse)와 같지 않다"라고 명시했다. 비관 시나리오는 이것이 반드시 일어날 것이라는 예측이 아니라, 전제 조건이 충족될 경우의 꼬리 위험에 대한 경고다.

논쟁의 핵심은 AI가 노동을 대체할 것인가가 아니라 어떤 속도로 대체할 것인가의 문제다. 역사적으로 기술 혁명은 언제나 새로운 일자리를 창출했다. 그러나 그 과정에서 기존 노동자들이 새로운 직종으로 이동하는 데 필요한 시간과 비용은 항상 불균등하게 분배됐다. 산업혁명 당시의 '엥겔스 멈춤(Engels Pause)'처럼 생산성 이익이 노동자에게 도달하기까지 수십 년의 격차가 존재했다.

Citrini 보고서의 진정한 가치는 파국 시나리오 그 자체에 있지 않다. 화이트칼라 노동 소득이 GDP 순환의 핵심 연결고리라는 구조적 사실과 그 연결고리가 끊어질 경우 전통적인 통화정책만으로는 대응이 불가능하다는 통찰에 있다. AI가 만드는 문제는 '유동성 부족'이 아니라 '기술적 대체'이기 때문이다.

그 외 주요 연구 결과

Goldman Sachs: AI가 3억 개 일자리에 영향 가능

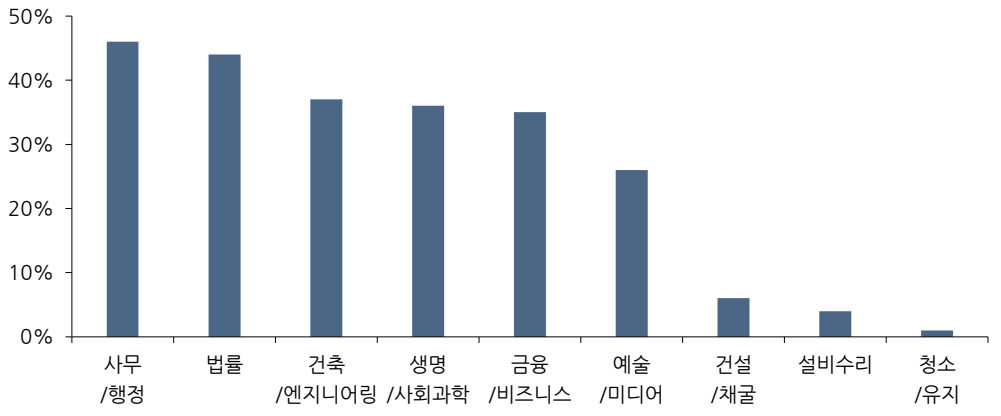
AI는 3억 개의 일자리 자동화로 노동 구조를 재편하는 동시에 생산성 혁신으로 글로벌 GDP를 7% 성장시킬 전망

AI가 노동시장에 미칠 영향을 가장 크게 제시한 연구 중 하나는 Goldman Sachs의 2023년 보고서이다. 해당 보고서는 생성형 AI 기술이 전 세계 노동시장에 미칠 잠재적 영향을 분석하며 약 3억 개의 정규직 일자리가 AI 자동화의 영향을 받을 수 있다고 전망했다.

보고서에 따르면 미국과 유럽 노동시장의 약 23% 수준의 직무(task)가 생성형 AI에 의해 자동화될 가능성이 있으며 일부 직무에서는 작업의 절반 이상이 AI로 수행될 수 있다. 특히 행정 업무, 법률 서비스, 금융 분석, 미디어 콘텐츠 제작 등 정보 기반 직무에서 자동화 가능성이 높게 나타났다. 또한 보고서는 AI 기술이 단순히 노동을 대체하는 것이 아니라 생산성 향상을 통해 장기적으로 글로벌 GDP를 약 7% 증가시킬 가능성도 제시했다. 이는 AI가 노동시장에 미치는 영향이 단순한 고용 감소가 아니라 생산성 증가와 노동 구조 변화가 동시에 나타나는 복합적 현상임을 의미한다.

특히 Goldman Sachs는 AI 자동화가 직무(task) 단위로 진행될 것이라고 강조한다. 즉 특정 직업이 완전히 사라지기보다는 해당 직업에서 수행되는 업무 중 일부가 AI로 대체되고 나머지 업무는 인간이 계속 수행하는 방식으로 노동 구조가 변화할 가능성이 높다는 것이다.

그림 25. 직종별 AI 자동화 가능 작업 비율



자료: Goldman Sachs, 상상인증권

IMF: 글로벌 노동자의 40%가 AI 영향

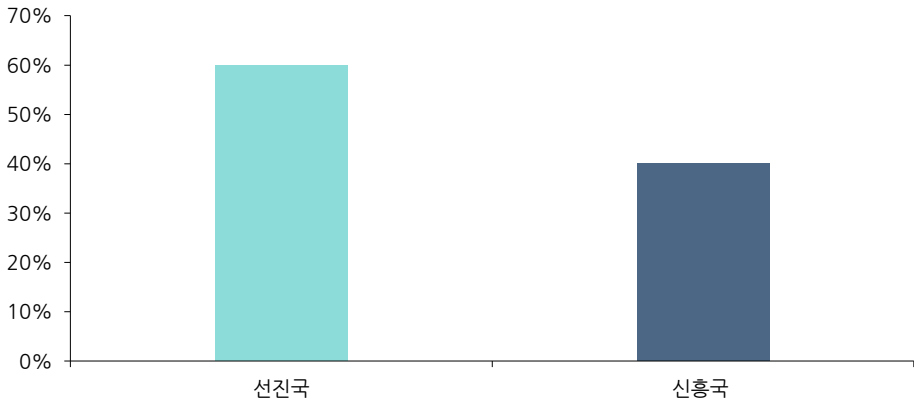
지식 노동 비중이 높은 선진국일수록 AI에 따른 자동화 위험과 소득 불평등이 심화

AI 기술이 노동시장에 미치는 영향은 선진국뿐 아니라 글로벌 노동시장에서도 중요한 이슈로 부상하고 있다. IMF는 2024년 보고서에서 전 세계 노동자의 약 40%가 AI 기술의 영향을 받을 것이라고 분석했다.

IMF 연구에 따르면 선진국에서는 약 60%의 직무가 AI 영향을 받을 가능성이 있으며 이 중 절반 정도는 자동화 위험이 존재하는 것으로 나타났다. 반면 신흥국에서는 AI 영향 비중이 약 40% 수준으로 분석되었다. 이러한 차이는 선진국 노동시장이 사무직과 지식 기반 직무 비중이 높기 때문으로 설명된다. 즉 제조업 중심 노동시장보다 오히려 서비스와 사무직 중심 노동시장에서 AI 영향이 더 크게 나타날 가능성이 있다는 것이다.

또한 IMF는 AI 기술이 노동시장 불평등을 확대할 가능성도 지적한다. AI 기술을 활용할 수 있는 고숙련 노동자에 대한 수요는 증가하는 반면 자동화 위험이 높은 직무에서는 임금과 고용 안정성이 악화될 수 있기 때문이다. 이는 결과적으로 소득 불평등 확대와 노동시장 양극화로 이어질 가능성이 있다.

그림 26. 선진국과 신흥국 간 직무가 AI에 영향을 받을 가능성



자료: IMF, 상상인증권

ESG 관점에서의 시사점

결론: AI 도입으로 인한 노사 갈등

AI가 노동시장에 도입되며 노동 대체와 노사 갈등을 확대하는 등 노동시장 구조 변화를 촉발

AI 기술의 확산은 노동시장 구조를 빠르게 변화시키고 있으며 그 영향은 단순한 자동화 수준을 넘어 노동의 성격과 노사 관계, 제도적 규제 환경까지 확장되고 있다. 특히 최근 미국과 유럽에서 나타난 사례들은 AI 도입이 기술적 문제를 넘어 사회적·제도적 갈등의 핵심 이슈로 부상하고 있음을 보여준다.

첫째, AI는 일부 직무를 자동화하면서 노동 대체 가능성에 대한 우려를 확대시키고 있다. 콘텐츠 검열, 고객 서비스, 운송 등 다양한 산업에서 AI 시스템이 인간 노동을 부분적으로 대체하기 시작하면서 고용 안정성에 대한 논쟁이 확대되고 있다. 이러한 변화는 기업의 생산성 향상과 비용 절감이라는 긍정적 효과를 가져오는 동시에 특정 직무의 축소나 구조조정 가능성을 높이며 노동시장 내 불확실성을 증가시킨다.

둘째, AI 확산은 노동의 '디지털 권리' 문제를 새롭게 제기하고 있다. 할리우드 배우 및 성우 노조의 파업 사례에서 나타난 것처럼 AI가 개인의 얼굴이나 음성을 학습해 디지털 복제물을 생성할 수 있게 되면서 노동자는 자신의 데이터와 정체성에 대한 권리 보호를 요구하고 있다. 이는 향후 노동 계약이나 산업 규범에서 AI 활용에 대한 동의, 보상, 사용 범위를 명확히 규정해야 한다는 새로운 제도적 요구로 이어지고 있는 상황이다.

셋째, AI는 노동 현장에서 알고리즘 기반 관리(algorithmic management)를 확산시키며 노동 통제 방식에도 변화를 가져오고 있다. 플랫폼 기업이나 디지털 산업에서는 AI가 노동자의 업무 배정, 성과 평가, 근무 관리 등을 수행하는 사례가 증가하고 있으며 이에 따라 노동조합은 인간 감독 의무나 알고리즘 투명성 확보와 같은 제도적 장치를 요구하고 있다. 실제로 유럽에서는 플랫폼 노동 지침을 통해 AI 의사결정에 대한 인간 감독을 강화하려는 정책 논의가 진행되고 있다.

이러한 변화 속에서 노조의 역할 또한 확대되고 있다. 미국과 유럽의 노동조합들은 단순히 자동화 반대를 넘어 AI 도입 과정에 노동자의 참여를 보장하고 노동권 보호를 제도화하려는 움직임을 보인다. 예를 들어 미국의 주요 노동단체들은 'Worker-Centered AI'와 같은 정책을 통해 AI 도입 시 노동자 참여, 데이터 권리 보호, 재교육 프로그램 확대 등을 요구하고 있다. 이는 AI 기술 발전이 노동시장에 미치는 영향을 완화하고 기술 발전의 혜택을 노동자와 공유하기 위한 정책적 대응으로 볼 수 있다.

표 5. AI 도입과 노동 갈등 사례

국가	산업	갈등 사례	노조 요구 및 법·정책 움직임
미국	영화·엔터테인먼트	2023년 배우 노조 파업	배우 얼굴·음성을 스캔해 AI로 재사용하는 것을 제한하고 배우 동의 및 보상 의무 요구
	게임	2024-2025 게임 성우 파업	AI가 성우 음성을 학습하거나 복제하는 것을 제한하고 AI 사용 시 계약·보상 규정 마련 요구
	자동차	자율주행 로봇택시 반대 시위	운전 노동조합이 자율주행 차량에 인간 안전 운전자 의무화 법안 지지
	전체	노동조합 정책 대응	대형 노조가 “Worker-Centered AI” 정책 발표, AI 도입 시 노동자 참여·재교육·감시 제한 요구
	항만·물류	자동화 장비 도입 갈등	항만 노조가 자동화 도입 시 고용 보장 및 임금 인상 요구
	교육	AI로 업무 대체 논란	노동조합이 AI로 파업 노동을 대체하는 행위에 반발
독일	플랫폼	AI 기반 알고리즘 관리 문제	TikTok이 AI 도입으로 150명 해고 계획 → 노조 ver.di 파업 및 보상 요구
유럽	플랫폼	콘텐츠 모더레이터 구조조정	EU 플랫폼 노동 지침 추진 → AI 의사결정의 인간 감독 의무

자료: 주요 언론보도 종합, 상상인증권

AI 활용과 그에 따른 ESG 영향분석

04

생성형 AI 확산에 따른 저작권 리스크 발생

AI의 저작권 이슈와 창작자 수익 악화

학습데이터 논란과 수익 배분 구조의 왜곡

학습데이터 출처 불투명성이 부각하는 저작권 문제

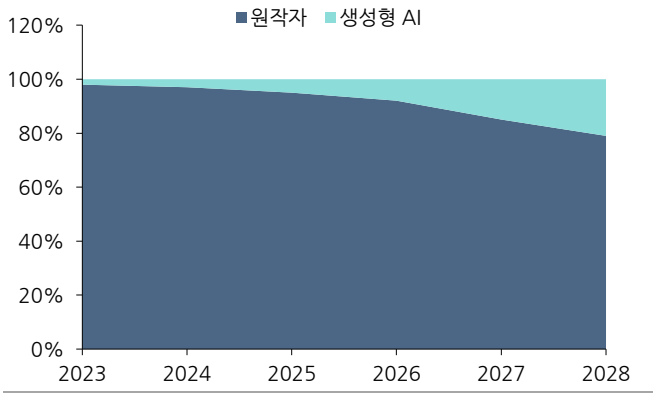
생성형 AI는 문서 작성 및 사진 제작 등 다양한 업무 영역에서 빠르게 확산되며 기업 생산성을 높이는 핵심 도구로 자리 잡고 있다. 그러나 이러한 혁신의 이면에는 AI 모델이 학습 과정에서 방대한 텍스트·이미지·음원 데이터를 어떤 방식으로 수집하고 활용했는지에 대한 불투명성이 존재한다. 특히 뉴스, 음악, 영상 등 저작권 보호를 받는 콘텐츠가 권리자 동의 없이 학습에 활용됐을 가능성은 이제 윤리 이슈를 넘어 실질적인 재무 리스크로 전환되고 있다.

음악·시청각분야 창작자 수익 감소 우려, 총 220억 유로 규모

국제저작권단체연맹(CISAC)은 2024년 연구에서 생성형 AI로 인해 2028년까지 음악 분야 창작자 수익의 24%, 시청각 분야 창작자 수익의 21%가 잠재적으로 위협받을 수 있다고 추산했다. 금액 기준으로는 향후 5년 누적 기준 음악 분야 약 100억 유로, 시청각 분야 약 120억 유로, 총 한화 약 38조 원 규모의 수익이 위험에 노출될 수 있다는 분석이다.

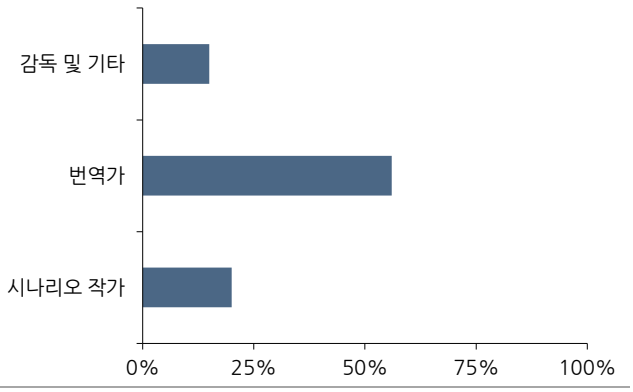
반면 같은 기간 생성형 AI 기반 음악 및 시청각 시장의 규모는 현재 약 30억 유로에서 2028년 640억 유로로 급증할 것으로 전망됐다. 즉, AI 사업자의 매출 확대와 원저작권자의 수익 잠식이 동시에 진행되는 구조라는 점에서, 데이터 무단 사용 이슈는 기술 발전의 부수적 논란이 아니라 가치 배분 구조의 왜곡으로 볼 필요가 있다. 이러한 우려는 이미 글로벌 언론사와 작가들을 중심으로 실제 소송으로 확산되고 있다.

그림 27. 생성형 AI에 따른 창작자 수익 배분 변화



자료: CISAC, 상상인증권

그림 28. 2028년 생성형 AI에 따른 창작자 유형별 수익 잠식률 전망



자료: CISAC, 상상인증권

생성형 AI 저작권 분쟁의 대표 사례

NYT vs Open AI

대표적인 분쟁 사례로는 뉴욕타임스(NYT)의 Open AI 제소가 있다. NYT는 Open AI가 자사 기사를 무단으로 학습한 뒤, 재현하거나 상세히 요약하는 방식으로 제공함으로써 저작권을 침해했다고 주장했다.

NYT가 특히 문제 삼은 부분은 이른바 '원문 재현(regurgitation)' 현상이다. NYT는 챗GPT가 유료 구독을 통해서만 접근 가능한 자사 기사와 리뷰를 장문 단위로 매우 유사하게 출력한 사례를 다수 제시했다. 이는 AI가 단순히 언어 패턴을 학습한 수준을 넘어, 원문 표현을 과도하게 재현하고 있다는 점에서 핵심 쟁점으로 부각됐다.

AI의 무단 학습이 뉴스 콘텐츠의 광고·구독 수익 잠식으로 이어질 가능성 지적

NYT는 이러한 행위가 단순한 권리 침해를 넘어 언론사의 수익모델 잠식으로 이어질 수 있다고 주장한다. 기존 검색 서비스는 기사 제목과 링크를 제공함으로써 이용자를 언론사 웹사이트로 유입시키는 구조였으나, 생성형 AI는 원문 방문 없이도 필요한 내용을 즉시 제공할 수 있다. 이에 따라 웹사이트 트래픽 감소, 디지털 광고 수익 하락, 신규 유료 구독자 유입 축소 등 직접적인 경제적 피해가 발생할 수 있다는 것이다.

또한 NYT는 이 사안이 정당한 콘텐츠 라이선스 시장을 훼손할 수 있다고 지적한다. 무단 학습이 선행될 경우, 향후 언론사가 데이터 제공에 대해 정당한 대가를 받을 유인이 약화될 수 있으며, 궁극적으로는 뉴스 콘텐츠 라이선스 시장 자체가 위축될 수 있다는 우려가 제기된다.

그림 29. NYT 소장의 '텍스트 비교(Regurgitation)'

Output from GPT-4:

exempted it from regulations, subsidized its operations and promoted its practices, records and interviews showed.

Their actions turned one of the best-known symbols of New York — its yellow cabs — into a financial trap for thousands of immigrant drivers. More than 950 have filed for bankruptcy, according to a Times analysis of court records, and many more struggle to stay afloat.

"Nobody wanted to upset the industry," said David Klahr, who from 2007 to 2016 held several management posts at the Taxi and Limousine Commission, the city agency that oversees medallions. "Nobody wanted to kill the golden goose."

New York City in particular failed the taxi industry, The Times found. Two former mayors, Rudolph W. Giuliani and Michael R. Bloomberg, placed political allies inside the Taxi and Limousine Commission and directed it to sell medallions to help them balance budgets and fund key initiatives.

During that period, much like in the mortgage lending crisis, a group of industry leaders enriched themselves by artificially inflating medallion prices. They encouraged medallion buyers to borrow as much as possible and ensnared them in interest-only loans

Actual text from NYTimes:

exempted it from regulations, subsidized its operations and promoted its practices, records and interviews showed.

Their actions turned one of the best-known symbols of New York — its signature yellow cabs — into a financial trap for thousands of immigrant drivers. More than 950 have filed for bankruptcy, according to a Times analysis of court records, and many more struggle to stay afloat.

"Nobody wanted to upset the industry," said David Klahr, who from 2007 to 2016 held several management posts at the Taxi and Limousine Commission, the city agency that oversees cabs. "Nobody wanted to kill the golden goose."

New York City in particular failed the taxi industry, The Times found. Two former mayors, Rudolph W. Giuliani and Michael R. Bloomberg, placed political allies inside the Taxi and Limousine Commission and directed it to sell medallions to help them balance budgets and fund priorities. Mayor Bill de Blasio continued the policies.

Under Mr. Bloomberg and Mr. de Blasio, the city made more than $855 million by selling taxi medallions and collecting taxes on private sales, according to the city.

자료: NYT, 상상인증권

Authors Guild(작가조합) vs Open AI

다른 대표적 사례로는 조지 R.R. 마틴을 포함한 미국 유명 작가들과 Authors Guild(작가조합)가 OpenAI를 상대로 제기한 소송이 있다. 원고들은 자신들의 베스트셀러 소설이 허가 없이 AI 학습에 활용되어 손해를 입혔다고 주장했다.

학습데이터 취득 경로와 저작권자 보상 체계가 핵심 문제로 부각

원고 측은 OpenAI가 방대한 도서 데이터를 확보하는 과정에서 정당한 구매나 라이선스 계약을 거치지 않았으며, 일부 불법 전자책 유통 사이트를 통해 저작물이 수집·활용됐을 가능성을 제기했다. 이는 단순한 무단 학습 여부를 넘어, 학습 데이터의 취득 경로 자체가 적법했는지를 둘러싼 핵심 쟁점으로 부각됐다.

작가들이 특히 문제 삼는 부분은 생성형 AI가 개별 작가의 문체, 서사 방식, 세계관의 특징을 높은 수준으로 모사할 수 있다는 점이다. 일례로, 이용자가 "조지 R.R. 마틴의 문체로 후속 이야기를 작성해 달라"는 식의 요청을 입력할 경우, 생성형 AI가 원작의 분위기와 문체적 특징을 연상시키는 결과물을 생성할 수 있다는 점이 반복적으로 지적돼 왔다. 원고 측은 이러한 출력물이 단순한 스타일 차용을 넘어, 원저작물의 창작성과 작가 고유의 표현 양식을 실질적으로 활용한 결과일 수 있다고 주장한다

또한 원고들은 이 같은 저작권 침해가 고의적 성격을 가진다고 보고, 침해 저작물 1건당 최대 15만 달러의 법정손해배상이 적용될 수 있다고 주장했다. 해당 사건은 생성형 AI의 성능 향상이 대규모 저작물의 무단 활용에 기반한 것인지, 그리고 AI 시대에 창작자의 보상청구권, 저작물 이용에 대한 동의권이 어디까지 보호되어야 하는지를 보여주는 상징적 분쟁으로 평가된다.

그림 30. 작가 조합과 오픈AI 저작권 분쟁

자료: The Guardian, 상상인증권

라이선스 확보를 통한 기업의 대응 움직임

콘텐츠 라이선스 계약 확대 사례

AI 기업 경쟁력, 모델 성능에서 적법한 데이터 확보 역량으로 확장

결국 생성형 AI 산업은 점차 정식 라이선스 확보 중심 전략으로 이동하는 흐름을 보이고 있다. 이는 저작권 분쟁이 확대되면서, AI 기업의 경쟁력이 단순한 모델 성능이 아니라 학습데이터의 출처 적법성, 권리자 보상 체계, 콘텐츠 활용에 대한 계약 역량까지 포함하는 방향으로 재정의되고 있음을 시사한다.

실제로 OpenAI는 News Corp와 5년간 2억5천만달러를 상회하는 계약을 체결했고, Condé Nast와도 다년 계약을 맺어 Vogue, The New Yorker, Wired 등 주요 콘텐츠를 자사 서비스에 활용할 수 있게 했다. 이 밖에도 AP통신, Axel Springer, TIME, Le Monde 등과의 제휴를 통해 고품질 뉴스 콘텐츠를 정식으로 확보하는 움직임을 확대해 왔다.

OpenAI, 언론·엔터테인먼트 콘텐츠 라이선스 확보 본격화

이 같은 변화는 텍스트 영역에만 국한되지 않는다. 2025년 12월 디즈니와 OpenAI는 3년 라이선스 계약을 체결하고, Sora와 ChatGPT Images에서 디즈니·스타워즈 등 일부 캐릭터와 관련 자산을 합법적으로 활용할 수 있도록 했다. 같은 발표에서 디즈니는 OpenAI에 10억달러를 투자한다고 밝혔는데, 이는 생성형 AI가 텍스트를 넘어 영상·이미지 영역으로 확장되는 과정에서도 무단 학습이 아니라 계약 기반 IP 활용 모델이 대안으로 자리 잡고 있음을 보여준다.

종합하면, AI 저작권 이슈는 더 이상 개별 기업의 소송 리스크에 그치지 않는다. 이는 콘텐츠 산업 전반의 수익 배분 구조와 창작 유인을 좌우하는 문제이자, AI 기업 입장에서는 향후 비용 구조와 사업 지속가능성을 결정짓는 핵심 거버넌스 이슈로 해석할 필요가 있다. 따라서 투자자 관점에서도 생성형 AI 기업의 경쟁력을 평가할 때는 모델 성능뿐 아니라, 학습데이터의 출처 투명성, 라이선스 체결 능력, 권리자 보상 체계, 분쟁 발생 시 대응 역량을 함께 점검할 필요가 있다.

표 6. AI 기업의 라이선스 계약 현황

AI 기업	콘텐츠/IP 보유사	추정 계약 규모	데이터 유형
OpenAI	뉴스 코프 (WSJ, Dow Jones 등)	약 3,400 억 원	언론 기사
	디즈니	약 1 조 3,500 억 원	캐릭터 IP 및 영상
Google	레딧	약 810 억 원	커뮤니티 게시글
Apple	셔터스톡	약 330 억 원	이미지, 오디오
Meta	뉴스 코프 (WSJ, Dow Jones 등)	약 670 억 원	언론 기사
	로이터 통신 (Reuters)	비공개	실시간 뉴스

자료: 주요 언론보도 종합, 상상인증권

Compliance Notice

- 본 자료에 기재된 내용들은 작성자 본인의 의견을 정확하게 반영하고 있으며 외부의 부당한 압력이나 간섭 없이 작성되었음을 확인합니다. (작성자:백영찬)
- 본 자료는 고객의 증권투자를 돕기 위한 정보제공을 목적으로 제작되었습니다. 본 자료에 수록된 내용은 당사 리서치센터가 신뢰할 만한 자료 및 정보를 바탕으로 작성한 것이나, 당사가 그 정확성이나 완전성을 보장할 수 없으므로 참고자료로만 활용하시기 바라며 유가증권 투자 시 투자자 자신의 판단과 책임하에 최종결정을 하시기 바랍니다. 따라서 본 자료는 어떠한 경우에도 고객의 증권투자 결과에 대한 법적 책임소재의 증빙자료로 사용될 수 없습니다.
- 본 자료는 당사의 저작물로서 모든 저작권은 당사에게 있으며 어떠한 경우에도 당사의 동의 없이 복제, 배포, 전송, 변형될 수 없습니다.
- 동 자료는 제공시점 현재 기관투자가 또는 제3자에게 사전 제공한 사실이 없습니다.
- 동 자료의 추천종목은 전일 기준 현재당사에서 1% 이상 보유하고 있지 않습니다.
- 동 자료의 추천종목은 전일 기준 현재 당사의 조사분석 담당자 및 그 배우자 등 관련자가 보유하고 있지 않습니다.
- 동 자료의 추천종목에 해당하는 회사는 당사와 계열회사 관계에 있지 않습니다.

상상인증권
sangsangin securities