Chapter **01**

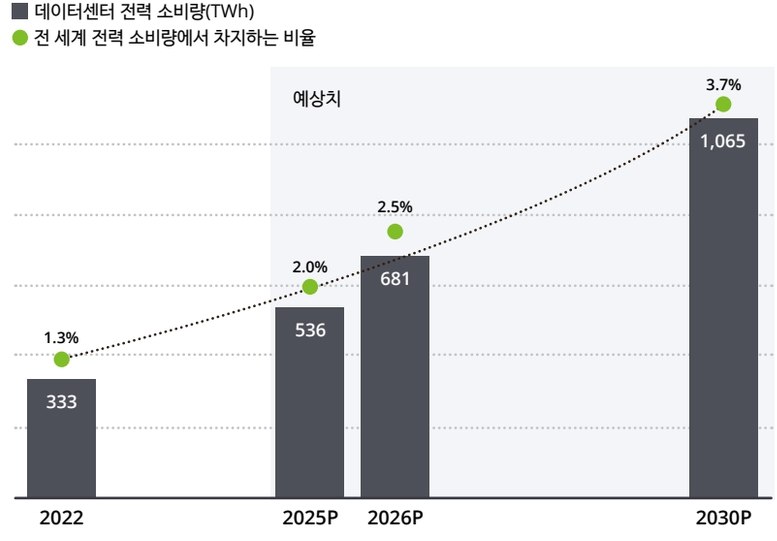
생성형AI 시대 본격화

01
생성형AI의 진화, 이제 데이터센터 에너지의 신뢰성과 청정화가 관건

> 첨단기술 산업은 데이터센터의 지속가능성을 개선하기 위해 인프라 최적화, 반도체칩 설계 재검토, 전력회사와의 협업 등 다각도의 방안을 모색할 필요가 있다.

인공지능AI 기반 데이터센터의 전력 소비가 계속 급증하고 있지만, 아직 전 세계 에너지 수요에서 데이터센터가 차지하는 비중은 크지 않다. 딜로이트는 2025년 데이터센터의 전력 소비량이 536테라와트시TWh로 전 세계 전력 소비량의 약 2%에 그칠 것으로 예상한다. 하지만 전력 소모가 큰 생성형AIgenerative AI 훈련과 추론 활동이 여타 활용사례나 앱에 비해 빠르게 증가하는 만큼, 2030년에 이르면 전 세계 데이터센터 전력 소비량이 약 1,065TWh에 달할 것으로 예상된다(그림 1).[1] 이처럼 급증하는 데이터센터의 전력 수요를 맞추면서 환경 영향을 줄이려면, 혁신적이고 에너지 효율적인 데이터센터 기술을 탐색함과 동시에 무탄소 에너지원의 활용을 늘려야 한다.

그림 1
전력 소모가 큰 생성형AI 모델을 중심으로 전 세계 데이터센터 전력 소비량 급증 전망



참조: 'P'는 예측값을 뜻함.
출처: Deloitte analysis based on publicly available information sources and conversations with industry experts.
방법론: 2022~2030년 전 세계 데이터센터 전력 소비량 추정치와 예상치는 미국 에너지정보청(EIA)이 주거·산업·운송 최종사용의 총 전력 사용량 데이터를 취합한 '국제 에너지 전망 2023'(International Energy Outlook 2023)의 기본 전력 소비 데이터에 기반해 집계했다. 전 세계 에너지 소비량에서 데이터센터가 차지하는 비율(%) 전망치는 세미 애널리시스(Semi Analysis), 미국 전력연구원(EPRI), 골드만삭스(Goldman Sachs), 블룸버그(Bloomberg), 래티튜드 미디어(Latitude Media)의 공개 자료와 더불어 첨단기술·에너지·지속가능성 분야 산업 전문가의 견해를 기반으로 도출했다.

물론 수많은 변수가 발생할 수 있으므로, 2030년과 그 이후의 전 세계 데이터센터 에너지 소비량을 정확히 예측하기는 쉽지 않다. 다만 AI와 데이터센터의 프로세싱 효율성이 지속적으로 개선되면 2030년에는 약 1,000TWh의 에너지가 소모될 것이라는 전망을 제시할 수 있다. 하지만

향후 수년 내로 이러한 효율성 개선이 이뤄지지 않는다면, 데이터센터의 전력 소비량은 1,300TWh를 돌파해, 전력회사들이 직접적 영향을 받고 기후중립 목표 달성에도 차질이 생길 수 있다.[2] 결과적으로 앞으로 10년간 AI 혁신에 박차를 가하고 데이터센터 효율성을 최적화해야만 지속 가능한 에너지 환경을 구축할 수 있다.

일부 지역에서는 이미 AI 데이터센터의 전력 수요 증가로 인해 전력 생산과 전력망 용량 관리가 해결하기 힘든 과제로 부각되고 있다.[3] 그래픽 처리장치GPU, 중앙처리장치CPU 서버, 스토리지 시스템, 냉각 시스템, 네트워크 스위치 등 데이터센터의 핵심 장치 및 시스템을 구동하는 데 필요한 전력이 2026년 글로벌 기준 96기가와트GW에 달해 2023년에 비해 약 두 배 증가할 것으로 예상된다. 이 중 AI 운영이 40% 이상을 차지할 것으로 보인다.[4] 전 세계 AI 데이터센터의 연간 전력 소비량은 2026년 90TWh에 달해 2022년에 비해 무려 10배 증가할 것으로 전망된다.[5] 이는 전 세계 데이터센터 2026년 전력 소비량 총합 전망치인 681Twh의 약 1/7 수준이다. 생성형AI 투자로 전력 수요가 급격히 촉발돼, 2024년 1분기 전 세계 AI 데이터센터의 전력 추가 순수요는 약 2GW로 2023년 4분기에 비해 25% 늘었고, 2023년 1분기에 비하면 세 배 이상 급증했다.[6] 더군다나 데이터센터 시설은 지리적으로 주로 미국에 집중돼 있어 전력 수요를 충당하기가 쉽지 않고, 일년 내내 24시간 운영되기 때문에 기존 전력 인프라에 부담을 주고 있다.[7]

이에 따라 첨단기술과 전력 산업이 힘을 합쳐 이러한 도전 과제를 해결하고 AI, 특히 생성형AI로 초래되는 에너지 부담을 줄이도록 노력해야 한다. 이미 다수의 빅테크 및 클라우드 기업들이 무탄소 에너지원에 투자하고 넷제로 목표를 향해 달리며 지속가능성 개선을 위한 노력을 펼치고 있다.[8]

하이퍼스케일러 업계, 고객 수요 증대에 발맞춰 생성형AI 데이터센터 대거 확대

이처럼 데이터센터 전력 수요가 급증하는 주요 원인은 전 세계적으로 하이퍼스케일러hyperscaler, 대규모 데이터센터 운용기업 업계가 데이터센터 용량을 확대하는 추세이기 때문이다.[9] 생성형AI를 중심으로 AI 수요가 증가할 것이라는 예상에 각국 정부와 기업들이 너도나도 데이터센터 확충에 나서고 있다. 특히 각국 정부는 기술 패권 싸움에서 지지 않기 위해 소버린 AIsovereign AI* 역량을 서둘러 구축하고 있다.[10] 전 세계 주요 하이퍼스케일러의 자본지출은 2024년 기준 약 2,000억 달러에서 2025년 2,200억 달러를 넘을 것으로 예상된다. 이를 기반으로 추산하면 전 세계 데이터센터 부동산 빌드아웃real estate build-out**은 이미 사상최대치를 찍었다.[11]

* 소버린 AI(sovereign AI)는 한 국가가 자체 데이터와 인프라를 기반으로 자국의 제도와 문화, 역사, 가치관을 심도 깊게 이해한 AI를 독립적으로 개발하고 운영하는 것을 뜻한다.
** 부동산 빌드아웃(real estate build-out)은 기본 구조만 갖춰진 상업용 또는 주거용 부동산을 임차인 및 사용자의 세부적인 요구에 맞춰 추가 공사나 마감 작업을 하는 것을 의미한다.

또한 딜로이트의 기업 대상 서베이에 따르면, 지금까지는 생성형AI 도입이 실험적 단계에 불과했지만,[12] 이를 통해 실질적 가치가 입증되면서 빠른 속도로 규모를 확대해 본격적 배치가 이뤄질 것으로 예상된다. 생성형AI 기술의 성숙과 함께 활용도 증가함에 따라, 하이퍼스케일러와 클라우드 기업들의 자본지출은 2025~2026년 높은 수준으로 유지될 전망이다.

데이터센터에서 전력을 가장 많이 소비하는 것은 컴퓨팅 파워computing power 및 서버 시스템(데이터센터 전력 소비의 약 40% 차지)과 냉각 시스템(약 38~40% 차지)이다. 이들은 AI 데이터센터에서도 가장 에너지 집약적 시스템이므로 지속적으로 데이터센터의 전력 소비를 촉발하는 요인이 될 것이다. 이 외 내부 전력 조정 시스템이 8~10%, 네트워크 및 통신 장비와 스토리지 시스템이 각각 약 5%, 조명 시설이 1~2%의 전력을 소비한다(그림 2).[13] 생성형AI로 막대한 용량의 전력이 필요한 만큼, 하이퍼스케일러와 데이터센터 운영 기업들은 데이터센터 설계 시 대체 에너지원, 새로운 형태의 냉각 시스템, 에너지 효율성 개선 솔루션 등을 모색할 필요가 있다.

그림 2

컴퓨팅파워와 냉각 시스템이 AI 데이터센터에서 가장 많은 전력 소비

출처: 딜로이트 분석

생성형AI로 데이터센터 전력 수요 급증

데이터센터의 에너지 소비는 AI 수요와 더불어 2023년부터 급증하기 시작했다.[14] 첨단 AI 시스템을 배치하려면 수많은 반도체칩과 처리 용량이 필요하고 복잡한 생성형AI 모델을 훈련하려면 수천 개의 GPU가 필요하다.

생성형AI와 고성능 컴퓨팅 환경을 운영하는 하이퍼스케일러와 대규모 데이터센터 기업들은 컴퓨팅 파워를 지원하기 위해 고밀도의 인프라가 필요하다. 과거 데이터센터는 반도체칩 한 개당 약 150~200와트W로 구동되는 CPU에 주로 의존했다.[15] AI용 GPU는 2022년까지 400W의 전력으로 가능했으나, 2023년 생성형AI용 최첨단 GPU는 700W가 필요했고, 2024년 차세대 반도체는 1,200W가 필요할 것으로 예상된다.[16] 현재 약 8종류가 개발된 차세대 반도체는 데이터센터의 서버 랙rack 내부에 10개씩 적재된 블레이드blade에 장착되는데, 불과 수년 전에 개발된 데이터센터 설계에 비해 기초면적 평방미터당 더 많은 물을 소비하고 더 많은 열을 발산한다.[17] 서버 랙당 전력 요구량 추정치는 2023년 36kW에서 2024년 초 20kW 수준으로 다소 줄었으나, 2027년 50kW로 급증할 것으로 예상된다.[18]

컴퓨터 성능을 나타내는 단위인 플롭스FLOPS, floating-point operations per second, 초당 부동소수점 연산를 기준으로 총 AI 컴퓨팅 용량 또한 생성형AI가 등

장한 이후 기하급수적으로 증가하고 있다. AI 컴퓨팅 플롭스는 2023년 1분기 이후 전 세계적으로 분기마다 50~60%씩 늘고 있으며, 2025년 1분기까지 같은 속도의 증가세가 예상된다.[19] 데이터센터의 컴퓨팅 용량은 플롭스뿐 아니라 MWh 및 TWh로도 측정할 수 있다.

수십억 개의 파라미터를 가진 생성형AI의 대규모 언어모델, 막대한 전력 소비

생성형AI의 대규모언어모델LLM은 갈수록 많은 파라미터parameter, 매개변수를 통합해 한층 고도화된다. 2021~2022년 1,000억~2,000억 개의 초기 파라미터 모델에서 시작해 2024년 중반 현재 첨단 LLM의 파라미터는 2조 개에 육박해 복잡한 이미지를 해석 및 해독한다.[20] 더 나아가 이제는 10조 개 파라미터를 갖춘 LLM 출시 경쟁에 불이 붙었다. AI 모델 훈련 및 배치를 위해 데이터 프로세싱과 컴퓨팅 파워에 더 많은 파라미터가 추가되며, 이는 생성형AI 프로세서와 액셀러레이터의 수요를 촉발해, 결국 전력 소비를 대거 끌어올린다.

게다가 LLM 훈련은 에너지 집약적 과정이다. 최근 조사에 따르면, 1,750억 개 이상의 파라미터로 훈련한 주요 LLM은 1회 훈련 시마다 324~1,287MWh의 전력을 소비하는 것으로 나타났다. 그리고 LLM은 재훈련이 필요한 경우가 많다.[21]

평균적으로 생성형AI 기반 프롬프트 요청을 처리하는 것은 인터넷 검색 결과 도출보다 10배에서 100배 많은 전력을 소비한다.[22] 딜로이트의 분석에 따르면, 일일 인터넷 검색 건수의 5%만 생성형AI 프롬프트 요청으로 바뀌어도, 개당 8개의 특수화된 GPU 코어가 탑재되어 평균 6.5kW의 전력을 소비하는 서버 약 2만 개가 필요하다. 일일 평균 전력 소비량으로 따지면 3.12GWh, 연 평균 소비량은 1.14TWh에 달하는 규모로,[23] 약 10만8,450 미국 가구의 연간 전력 소비량과 맞먹는 수준이다.[24]

데이터센터의 전력 수요, 전력 산업 전환의 과제이자 기회

전력 산업은 이미 수요 증가에 대비하며, 일부 국가에서는 2050년까지 전력 소비량이 최대 세 배 증가할 것으로 예측하고 있다.[25] 하지만 최근 데이터센터의 전력 수요가 급증하면서 일부 지역에서는 전력 소비량 증가세도 가속화되고 있다. 상당수 국가에서 기존의 전력 수요량 전망치는 전동화를 중심으로 데이터센터의 전력 소비 및 전반적인 경제성장을 기반으로 도출됐다. 하지만 지금 가시화되는 데이터센터 전력 소비 급증 양상은 빙산의 일각에 불과하며, 향후 전력 산업 전반의 전환을 초래할 정도로 가속화될 수 있다.[26]

현재 전력 산업은 수십년에 걸친 전환을 아직 진행 중이다. 전력망 인프라를 구축·업그레이드·탈탄소화하고, 시스템과 자산을 디지털화하며,

갈수록 극심해지는 기후 환경에서도 피해를 최소화하기 위해 설비 방재를 강화하고, 증가하는 사이버보안 위협에 대응해 네트워크 보호 시스템을 구축하고 있다.[27] 하지만 여전히 일부 국가에서는 전력망이 수요를 따라잡지 못하고 있으며, 특히 저탄소 및 무탄소 전력은 공급이 턱없이 부족하다. 미국의 경우 2026년 전국 전력 총소비량에서 데이터센터가 6%(260TWh)를 차지할 것으로 예상된다.[28] 영국은 AI 데이터센터를 중심으로 데이터센터 전력 수요량이 단 10년 만에 6배 증가할 것으로 전망된다.[29] 중국도 2026년 전력 총수요량에서 AI를 포함한 데이터센터가 6%를 차지할 것으로 예상된다.[30] 중국의 경우 주요 전력원이 여전히 석탄이기 때문에 데이터센터가 대기오염의 또 다른 원인이 될 수 있다. 2021년 기준 중국에서 석탄이 에너지 사용의 61%, 이산화탄소 배출의 79%를 차지했다.[31]

일부 국가들은 데이터센터의 전력 수요 증가를 억제하기 위해 규제 대응에 나섰다. 아일랜드는 기존 데이터센터들이 전력 총소비량의 1/5을 차지하는 가운데 AI 데이터센터가 늘면서 이에 따른 전력 소비도 한층 늘어날 전망이다.[32] 가계 전력 소비량은 오히려 줄어들고 있음에도 전력 공급난 가능성이 대두되자, 아일랜드 정부는 한시적으로 신규 데이터센터의 전력망 연결을 금지하기도 했다.[33] 네덜란드 암스테르담 시정부도 지속 가능 도시 개발을 위해 데이터센터 신설을 금지했다.[34] 싱가포르는 데이터센터 내 온도를 26℃ 이상으로 유지하도록 하는 새로운 지속가능성 표준을 도입했다. 데이터센터 내 유지 온도가 높아지면 냉각 시스템 가동이 줄어 전력 소비량은 감소하지만, 반도체칩의 수명이 짧아진다는 문제가 있다.[35]

데이터센터는 지리적으로 집중돼 있는 데다 전력 수요를 당장 365일 24시간 충족할 필요가 있어, 그렇지 않아도 전동화와 제조업 전환 등으로 발생하는 새로운 전력 수요 부담을 안고 있는 첨단기술 및 전력 기업들이 한층 복잡한 과제를 마주하게 됐다. 전 세계에서 가장 큰 데이터센터 시장은 미국 노던 버지니아Northern Virginia에 형성돼 있는데,[36] 현지 전력 및 유틸리티 회사인 도미니언 에너지Dominion Energy는 향후 15년 새 데이터센터의 전력 수요가 네 배 증가하면서 전력 총수요가 약 85% 증가할 것으로 내다봤다.[37] 하지만 많은 테크 기업들이 원하는 365일 24시간 무탄소 전력 공급은 단기 내에 실현하기가 어렵다. 이에 전력회사들은 늘어나는 전력 수요를 충당하면서도 공급 신뢰성과 경제성을 사수하기 위해 다각도로 모색하고 있다. 새로운 재생에너지와 배터리 저장 시설의 개발 및 구축 외에도, 급한 불을 끄기 위해 천연가스 발전소 신설 계획도 세우고 있다.[38] 천연가스는 석탄보다는 양호하지만 탄소를 배출하는 에너지원이기 때문에, 이로 인해 전력회사와 주 정부, 더 나아가 국가의 탈탄소화 목표 달성이 한층 어려워질 수 있다.[39]

다만 막대한 규모의 에너지를 소비하는 AI가 청정에너지 전환을 가속화하는 요인으로 작용할 수도 있다. 일부 유틸리티 회사들은 이미 AI를 활용해 날씨 및 부하 예측의 정확도를 높이고, 전력망과 재생에너지 자산

성능을 강화하고, 폭풍 재해의 복구 속도를 끌어올리고, 산불 위험 예측을 개선하는 등 전력망 운영의 비용을 절감하고 효율성과 신뢰성을 개선하고 있다.[40]

물 자원 소모가 큰 데이터센터 냉각 시스템

차세대 CPU와 GPU는 이전 세대보다 열 밀도thermal density*가 높은 성질이 있다. 또한 일부 서버 벤더들은 고성능 컴퓨팅과 AI 앱의 증가하는 수요를 맞추기 위해 서버 랙에 전력을 많이 소비하는 반도체칩을 더 많이 적재하고 있다. 하지만 생성형AI 등 서버 랙의 밀도가 높아질수록 냉각수가 더 많이 필요하다. AI 데이터센터의 민물 수요량은 2027년 최대 1.7조 갤런에 달할 것으로 전망된다.[41] 만약 에어 냉각air-based cooling이나 증발 식수로 초과 열을 관리하면 매년 5,000만 갤런 이상의 물이 필요하다. 이는 약 1만4,700개의 스마트폰을 만드는 데 필요한 양이다.[42] 이렇게 사용된 물은 대수층이나 저수지, 상수도 등으로 되돌릴 수도 없다.[43]

* 열 밀도(thermal density)는 데이터센터나 에너지 시스템 등 특정 공간 내에서 발생하는 열 에너지의 양을 의미하며, 열 관리의 효율성을 측정하는 지표로 사용된다.

에어 냉각 시스템만 가동해도 일반적 데이터센터 전력 소비량의 최대 40%가 소요된다. 따라서 데이터센터들은 액체 냉각 등 기존 에어 냉각 방식을 대체할 수 있는 방안을 모색해야 한다. 액체 냉각은 에어 냉각보다 열전사 성질이 강해 고밀도 서버 랙 냉각 시 에어 냉각보다 전력 소비량을 최대 90% 절감할 수 있다.[44] 액체 냉각은 서버 랙을 직접 냉각하기 때문에, 반도체칩으로 밀도가 높아진 서버 랙이라 하더라도 50~100kW의 정도의 전력만 소요해도 냉각이 가능하다.[45] 또한 냉각수 생산에 일반적으로 사용되는 냉각 장치도 필요하지 않다.

액체 냉각은 이처럼 데이터센터의 에너지 절감에 큰 도움이 되지만,[46] 아직 발전 초기 단계여서 전 세계적으로 광범위하게 AI 데이터센터에 도입되지 못하고 있다.[47] 게다가 물은 유한한 자원이기 때문에, 비용과 접근성이 어떻게 개선되느냐가 향후 광범위한 상용화 여부를 결정지을 것으로 예상된다.

지속 가능 솔루션과 무탄소 에너지원 모색하는 테크 산업

빅테크 기업들은 AI 데이터센터의 무탄소 전력 공급을 가속화하기 위해 에너지 기업들과 전력 매입계약 및 장기 계약을 맺는 방식으로 재생에너지 확대 방안을 적극적으로 모색하고 있다.[48] 이러한 계약에 힘입어 재생에너지 프로젝트를 추진하는 주체들은 필요한 자금을 확보할 수 있다. 일부 테크 기업들은 전력 회사 및 혁신 기업들과 손잡고 첨단 지열 발전, 첨단 풍력 및 태양광 기술, 수력 발전, 수중 데이터센터 등 유력한 에너지 기술을 실험하고 규모를 확대하는 노력을 펼치고 있다.

일부 지역에서는 전력망 제약과 상호연결 기간 장기화 등으로 인해 새로운 재생에너지와 배터리 저장 시설을 전력망에 연결하는 작업이 지연되고 있다.[49] 수요는 많은데 전송 인프라가 불충분해 미국의 경우 관련 작업이 최대 5년 지연될 수 있다. 이에 따라 온사이트onsite 또는 독립형 off-grid 에너지 솔루션을 모색하거나,[50] 장기 에너지 저장LDES, long-duration energy storage이나 소형 모듈 원자로SMR, small modular nuclear reactor와 같은 신기술에 투자하는 테크 기업들이 늘고 있다. 테크 기업들과 유틸리티 회사들이 혁신적인 청정 에너지 기술의 규모를 확대하기 위해 협업을 계획하는 경우도 있다. 이러한 움직임은 궁극적으로 여타 조직의 청정 에너지 활용에도 도움이 되고 전력망의 탈탄소화를 앞당기는 효과를 가져올 수 있다.[51] 하지만 이러한 연구개발R&D 프로그램과 시범도입, 여타 청정에너지 투자 중 상당수는 수년 후에나 실질적 가치를 창출하고 투자수익을 증명하며 상업성이 가시화될 것이다.[52] SMR의 경우 아직 개발 초기 단계여서 단기 내 무탄소 솔루션으로 활용하기가 어렵다.[53]

미국 기업들 중에서 재생에너지를 가장 적극 도입하는 것은 테크 부문이다. 2024년 2월 28일까지 12개월 간 약 200건의 재생에너지 매입 계약 규모를 추적한 결과, 테크 기업들이 68% 이상을 조달한 것으로 나타났다.[54] 인도에서는 하이퍼스케일러와 데이터센터 운영 기업들이 태양광 도입을 주도하고 있다.[55] 테크 부문의 이 같은 움직임이 없다면, 상당수 재생에너지 프로젝트는 무산될 수 있다.[56]

이처럼 청정에너지 기술의 규모 확대를 위해 테크 부문의 자본 지원이 앞으로도 중요한 요인으로 작용할 전망이다. 테크 기업들은 혁신 기업이나 재생에너지 기업들과 직접 협업하기도 하고 유틸리티 기업들과 파트너십을 맺기도 한다.[57] 혁신 기업이나 전력 산업은 테크 부문보다 자금력이 현저하게 떨어지기 때문에, 테크 산업에서 유입되는 이러한 자본이 청정에너지 전환을 위해 매우 중요한 역할을 하고 있다.

결론

생성형AI 범용화에 따라 증가하는 데이터센터의 전력 수요를 지속적으로 충당하기 위해 테크 산업 전반과 하이퍼스케일러, 데이터센터 운영기업, 유틸리티 회사, 규제당국들은 어떠한 노력을 기울여야 할까? 데이터센터의 에너지 이슈는 딜로이트가 2021년 분석한 클라우드 이전과 비슷한 양상을 보일 가능성이 있다.[58] 수요 촉발 요인도 다르고 변화의 속도도 한층 가팔라졌지만, 테크 산업이 지속가능성과 AI 모델 출시의 시간 경쟁 사이 균형을 잡아야 하는 과제는 동일하다. 이와 동시에 데이터센터의 에너지 수요 증가를 억제하면서 생성형AI 등 AI 모델을 위한 전력을 더욱 지속 가능한 방법으로 공급할 방법을 모색해야 한다.

1. 생성형AI 반도체칩의 에너지 효율성 개선

차세대 AI 반도체는 이미 90일만에 AI 훈련을 수행할 수 있다. 이에 필

요한 전력은 8.6GWh로, 이전 세대 반도체칩이 동일한 데이터로 동일한 기능을 수행하는 데 필요한 에너지의 1/10도 되지 않는다.[59] 반도체 회사들은 전반적인 반도체 생태계와 지속적으로 협업해 와트당 플롭스FLOPS/watt를 개선해 훨씬 적은 전력으로 지금보다 훨씬 많이 AI 훈련을 수행할 수 있는 반도체칩을 개발하는 데 주력해야 한다.

2. 생성형AI 활용 최적화 및 에지edge 기기의 프로세싱 기능 확대

AI 모델의 훈련과 추론을 데이터센터 또는 에지 기기 중 어디에서 수행하는 것이 에너지 효율적인지 판단해, 데이터센터 설비 니즈의 균형을 맞출 필요가 있다. 응답 시간이 매우 중요한 태스크를 수행하거나 민감한 데이터를 다루고 엄격한 개인정보보호가 필요한 경우 에지 기기를 선택하는 것이 나을 수 있다. 생성형AI 작업부하 중 선별된 것만 데이터센터에서 처리하고 나머지는 로컬 장치나 근거리near-location 및 코로케이션co-location 장치에서 처리하면 네트워크 및 서버 대역폭의 부담도 줄일 수 있다.[60]

3. 생성형AI 알고리즘 수정 및 AI 작업부하 조정

AI 기초 모델은 규모 경쟁이 불필요할 수도 있다. 파라미터가 1조 개가 넘는 AI 기초 모델 대신 규모가 더 작은 모델만으로도 충분하고 오히려 지속가능성에 훨씬 도움이 될 수도 있다. 일부 스타트업들은 이미 클라우드 상의 에너지 집약적 컴퓨팅이 필요하지 않은 온디바이스on-device 멀티모달multimodal AI 모델을 개발하고 있다.[61] 이러한 모델을 사용하는 고객들은 AI 작업부하의 규모를 세부 조정해 실제 사업 니즈에 맞춰 기존 모델을 활용하고 필요할 때만 훈련을 수행해 맞춤형 생성형AI 모델을 활용할 수 있다. 이러한 방식으로 에너지 사용을 최소화할 수 있다. 또한 절대 지연되어서는 안 되는 실시간 추론이 필요한 특별한 경우 CPU가 비용과 사용 측면에서 훨씬 효율적일 수 있다.[62]

4. 전략적 파트너십을 활용해 지역별 또는 클러스터형 AI 데이터센터 니즈 충족

대학교와 같이 생성형AI 데이터센터 역량을 자체적으로 갖추기가 어려운 중소 규모 조직들은 데이터센터 운영 기업이나 클라우드 서비스 기업들과 파트너십을 맺어 문제를 해결할 수 있다. 이러한 파트너십을 통해 규모가 작은 고성능 컴퓨팅 GPU 클러스터 코로케이션에 고성능 컴퓨팅 솔루션을 제공할 수 있다.[63] 데이터센터 운영 기업들 입장에서는 사용 현황을 동적으로 추적함으로써 잠재적 기회와 수요 현황을 파악해 즉각 코로케이션 서비스에 적용할 수 있다는 이점이 있다.

5. 여타 이해당사자 및 산업 부문과 협력해 긍정적 환경 영향 창출

하이퍼스케일러와 고객들, 외부 데이터센터 운영 기업, 코로케이션 서비스 기업, 전력회사, 현지 규제당국 및 정부, 부동산 기업들까지 다양한 생태계 참가자들이 비즈니스와 환경, 사회를 위해 실행 가능한 부분이 무엇인지에 대한 논의를 지속해야 한다.[64] 데이터센터 기업이 컴퓨팅 핑 서버 자원을 몇몇 기업에 임대할 경우 발생하는 전략적 코로케이션 니즈 가

능성을 파악하는 것부터, 액체 냉각 시스템 내 적정 온도 등 냉각 니즈를 측정하고, 열과 폐수 관리 솔루션을 파악하고, 재활용 니즈를 계산하는 것까지 다양한 측면을 아우르는 협력이 필요하다. 일례로 유럽에서는 데이터센터에서 발생하는 폐열이 인근 수영장 물을 데우는 데 사용된다.[65] 전력회사들은 테크 산업과 더욱 긴밀히 협력해 데이터센터의 에너지 수요를 충족할 방법을 모색함과 동시에 이러한 파트너십을 통해 신에너지 기술 개발과 규모 확대를 위한 자금 확보 방안도 탐색해야 한다. 이러한 노력은 청정에너지의 전력망 통합 확대를 위해 매우 중요하다.

하이퍼스케일러와 전력 부문이 손을 잡고 생성형AI 데이터센터에 초점을 맞춰 무탄소 에너지원의 사용을 확대하려는 노력을 펼치면 분명 장기적으로 실질적 효과가 나타날 것이다.

02
생성형AI, 여성 소비자와 AI 인력 기반 확대 필요

> 여성들이 생성형AIgenerative AI의 혜택을 100% 누리려면, 테크 기업들이 AI 모델의 신뢰성을 강화하고 편향성을 축소하며 여성 AI 인력의 비중을 늘리기 위해 부단한 노력을 기울여야 한다.

여성의 실험적 또는 실제적 생성형AI 활용이 2025년 말에 이르면 미국에서 남성을 추월할 것으로 전망된다.[1] 2023년만 해도 여성의 생성형AI 활용은 남성의 절반에 그쳤으나, 여성의 활용이 빠른 속도로 증가하고 있다.[2] 이러한 변화 양상은 전 세계에서 공통적으로 나타나고 있다. 유럽에서 생성형AI 활용에 대한 서베이를 실시한 결과, 성별 격차가 여전히 컸으나 미국과 마찬가지로 여성의 활용이 빠르게 증가해[3] 앞으로 2년 내 성별 격차가 거의 좁혀질 것으로 예상된다. 생성형AI 활용의 성별 격차를 둘러싼 과제와 기회 또한 전 세계가 미국과 비슷한 양상을 보이고 있다.

생성형AI가 빠르게 범용화되고 있지만, 생성형AI를 제공하는 기업들의 데이터 보안 방식을 신뢰하지 못한다는 비율은 여성이 남성보다 높게 나타났다.[4] 이처럼 기술 신뢰의 성별 격차가 지속되면, 생성형AI를 일상적으로 사용하고 새로운 생성형AI 앱을 적극 사용하는 여성이 남성에 비

해 훨씬 적을 것이고 이는 향후 여성의 생성형AI 제품 및 서비스 구매도 위축시킬 수 있다. 테크 기업들은 이러한 신뢰의 격차를 극복하기 위해 데이터 보안을 강화하고 보다 투명한 데이터 관리 정책을 이행하고, 보다 광범위한 데이터 관리를 실행해야 한다.

AI 모델의 편향성 또한 신뢰에 부정적 영향을 준다.[5] AI 인력에서 여성이 차지하는 비중은 1/3이 채 되지 않으며,[6] AI 부문 근로자 대부분은 AI 산업이 남성 위주로 지속되는 한 AI 모델의 성별 편향성이 계속될 수 있다고 지적했다.[7] 따라서 AI 부문의 여성 인력을 늘려야 AI 모델의 편향성을 줄이고 관련 기술의 미래 방향을 결정하는 과정에서 여성이 더욱 큰 역할을 할 수 있다.

생성형AI 활용의 성별 격차, 빠르게 축소

최근 딜로이트 조사에 따르면, 다양한 지역에서 생성형AI 도입에 성별 격차가 있는 것으로 나타났다. 딜로이트가 지난 2년간 소비자들을 대상으로 디지털 라이프의 일환으로서 생성형AI 활용을 조사한 결과,[8] 미국 여성은 남성에 비해 생성형AI 활용에 있어 뒤처지는 것으로 나타났다(그림 1). 2023년 생성형AI를 시험 삼아서 사용하는 것을 넘어 프로젝트나 업무를 위해 사용한다는 비율은 여성(11%)이 남성(20%)의 약 절반에 그쳤다. 2024년 서베이에서 전반적인 사용 비율은 두 배 이상 뛰었으나, 여성(33%)과 남성(44%) 간 격차는 여전했다.

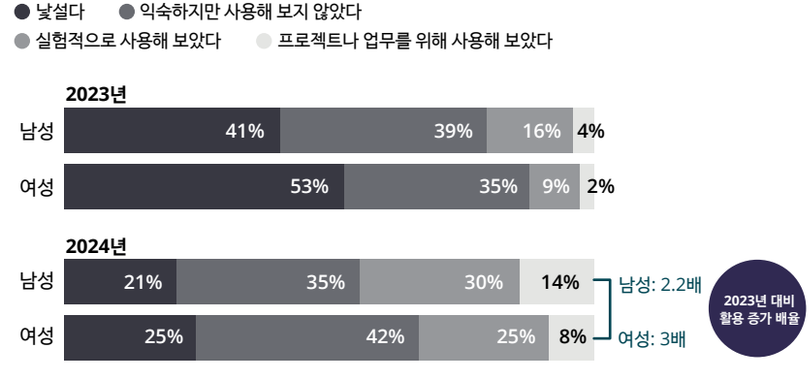
영국 소비자 대상 서베이에서도 비슷한 성별 격차가 나타났다. 2024년 서베이에서 생성형AI를 실질적으로 사용한다는 비율이 여성 28%, 남성 43%를 기록했다.[9] 12개 유럽국 소비자 대상 서베이에서는 두 자릿수 성별 격차가 나타났다.[10]

하지만 미국에서 여성이 남성을 빠르게 추격하고 있다. 2024년 생성형AI를 실질적 목적으로 사용했다는 여성의 비율이 세 배 늘어, 남성의 증가 배율 2.2배를 능가했다.[11] 이러한 추세가 지속되면 2025년 말에는 여성의 비율이 남성을 완전히 추격하거나 능가할 것으로 예상된다.[12]

그림 1

생성형AI 활용의 성별 격차가 여전히 존재하지만 빠르게 축소되는 중

2023~2024년 생성형AI 시스템 및 기술에 대한 인식 조사



표본 수: 미국 여성 소비자 1,040명(2023년)/1,992명(2024년);
미국 남성 소비자 962명(2023년)/1,841명(2024년)
출처: Deloitte 2024 Connected Consumer Survey, 5th edition; Deloitte 2023 Connected Consumer Survey, 4소 edition.

생성형AI에 대한 여성의 상대적 신뢰 부족, 관련 제품의 구매 위축으로 이어진다

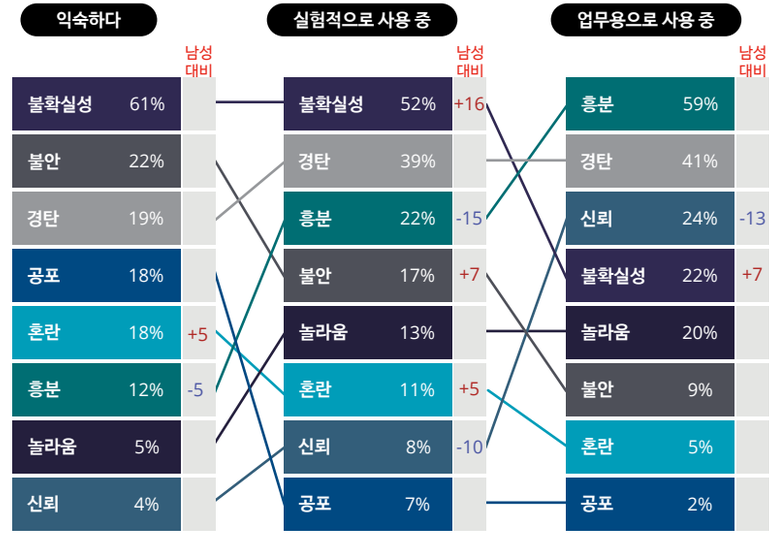
생성형AI 활용의 성별 격차가 좁혀지는 것은 고무적이지만, 그렇다고 여성들이 일상 업무에 생성형AI를 남성만큼 많이 사용한다는 의미는 아니다. 딜로이트의 2024년 서베이에 따르면, 생성형AI 사용자 중 하루에 한 번 이상 생성형AI를 사용한다는 비율이 여성 34%로 남성의 43%에 비해 여전히 낮았다.[13] 또 업무용으로 생성형AI를 사용한다는 응답자 중 생성형AI로 생산성이 대폭 개선됐다는 비율도 여성이 41%로 남성의 61%보다 낮았다.[14] 테크 기업들과 생성형AI 도입으로 효과를 얻으려는 기업들은 이러한 성별 격차를 유념하고 여성의 활용을 확대하기 위해 적극적인 노력을 기울일 필요가 있다.

이 같은 성별 격차가 나타나는 일부 원인은 신뢰에 대한 인식에 큰 간극이 있기 때문이다.[15] 생성형AI에 대한 여성의 인식이 '기술에 대한 익숙함'에서 '실제 실험적으로 사용해볼 수 있다'로 옮겨가면서, 불확실성과 불안, 공포, 혼란 등 부정적 감정은 줄고 경탄, 흥분, 놀라움, 신뢰 등 긍정적 감정은 증대하고 있다(그림 2).[16] 하지만 실험적 사용과 업무에 실제 활용하는 경우 모두 여성은 남성보다 생성형AI에 대한 신뢰가 낮았고 불확실성의 감정은 여전히 더 높았다. 생성형AI를 사용한다는 응답자 중 생성형AI 기업이 개인정보를 안전하게 보호할 것이라는 데 '높은 신뢰' 또는 '매우 높은 신뢰'를 가지고 있다고 답한 비율은 여성이 18%에 그쳐 남성의 31%에 비해 매우 낮게 나타났다.[17]

그림 2

생성형AI에 대한 여성의 부정적 감정은 줄고 긍정적 감정은 늘고 있지만, 신뢰는 여전히 남성보다 낮은 수준

생성형AI에 대해 '익숙하다' '실험적으로 사용하고 있다' '업무용으로 사용하고 있다'고 답한 여성들이 생성형AI에 대해 가장 크게 느끼고 있는 두 가지 감정을 선택하게 한 뒤 응답자 비율을 집계



표본 수: 생성형AI에 익숙하다는 미국 소비자=여성 827명/남성 647명; 생성형AI를 실험적으로 사용 중이라는 미국 소비자=여성 496명/남성 547명; 생성형AI를 업무용으로 사용 중이라는 미국 소비자=여성 170명/남성 259명
참조: 남성 대비 수치는 5%포인트 이상 격차일 때만 표시
출처: Deloitte 2024 Connected Consumer Survey, 5th edition.

신뢰 격차는 생성형AI에만 국한된 것은 아니며 전반적인 테크 서비스 및 상호작용에서도 나타난다. 딜로이트 2024년 서베이 결과, 온라인 서비스로 얻는 이점이 개인정보보호 침해 우려를 상쇄할 만큼 많다는 여성

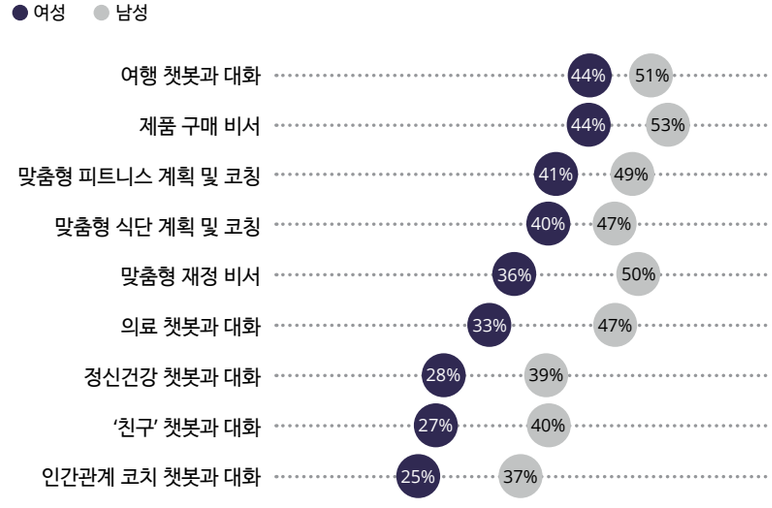
비율이 54%로 2023년의 46%에서 상승했음에도 불구하고, 62%를 기록한 남성보다 낮았다.[18] 딜로이트 보고서에 따르면, 여성은 개인정보의 활용 및 보호 방식에 대해 남성보다 우려가 높고 이로 인해 민감한 건강 및 피트니스 정보 등 개인정보를 공유할 의향도 훨씬 낮은 것으로 나타났다.[19] 여성은 보안 침해나 데이터 남용의 잠재적 피해를 남성보다 심각하게 인식하기 때문으로 풀이된다.[20]

생성형AI의 범용화가 심화될수록 개인정보보호와 데이터 보안을 둘러싼 뿌리깊은 우려가 다시금 부각될 수 있다.[21] 사용자가 생성형AI와 소통할 때, 시스템은 사용자의 데이터를 AI 모델에 입력할 수 있는데, 전문가들은 개인의 데이터가 AI 훈련에 활용되지 않도록 선택하기가 쉽지 않고 그 방식도 분명하지 않다고 지적했다.[22] 소비자들이 생성형AI와 민감하고 개인적인 사안에 대해 소통을 시작하는 만큼, 개인정보보호와 데이터 보안이 침해됐을 때의 피해도 그만큼 클 수 있다. 뿐만 아니라 개인정보보호와 데이터 보안의 신뢰와 관련해 성별 격차가 지속됨에 따라, 다양한 생성형AI 경험에 대한 관심에서도 격차가 나타나고 있다(그림 3).[23] 딜로이트 서베이 결과, 여행·쇼핑·피트니스·식단 등 덜 민감한 주제에 대해서조차 생성형AI와 소통하는 것에 여성이 남성보다 관심이 적은 편으로 나타났는데, 개인재정·인간관계·신체건강·정신건강 등 더욱 민감한 주제에 대해 소통하는 것은 여성이 남성보다 관심이 '현저히' 떨어지는 것으로 나타났다.

그림 3

여성은 다양한 생성형AI 경험에 대한 관심이 남성보다 낮은 수준

다음의 생성형AI 경험에 대해 '관심이 높다' 또는 '관심이 있다'고 답한 응답자 비율



표본 수: 생성형AI에 익숙하거나, 실험적으로 사용 중이거나, 업무용으로 사용 중이라는 미국 소비자=여성 1,492명/남성 1,454명

출처: Deloitte 2024 Connected Consumer Survey, 5th edition.

신뢰의 성별 격차는 새로운 생성형AI 기술의 구매 양상에서도 차이를 만든다. 이제 테크 기업들은 AI 칩을 탑재해 실시간으로 정보를 요약하고 사진과 영상을 생성하고 타 언어간 즉시 번역도 해내는 임베디드 AIembedded AI 노트북과 태블릿, 스마트폰을 출시하기 시작했다.[24] 하지만 딜로이트 서베이 결과 AI 기능이 추가되는 것이 기기 업그레이드 계획에 영향을 미친다는 여성의 비율이 남성보다 낮았다.[25] 구체적으로, 임

베디드 AI 스마트폰이 출시되면 계획보다 빨리 기기를 업그레이드할 것이라는 남성은 43%에 달했지만, 여성은 32%에 그쳤다. 반대로 임베디드 AI 스마트폰 출시에도 기기 업그레이드 계획을 전혀 바꾸지 않겠다는 비율은 여성이 58%로 50%를 기록한 남성보다 높게 나타났다. 임베디드 AI 노트북의 경우도 마찬가지로 업그레이드 계획을 앞당길 것이라는 여성은 28%에 불과해 41%를 기록한 남성보다 낮았다. 소비지출의 약 85%(추정치)를 여성이 결정하거나 영향력을 행사하는 만큼, 테크 기업들은 AI 기기 업그레이드에 대해 여성의 관심이 떨어진다는 점을 유념할 필요가 있다.[26]

여성의 생성형AI 활용 극대화를 가로막는 장애물은 신뢰 문제만이 아니다. 딜로이트 서베이 결과 소속 회사가 업무 시 생성형AI의 활용을 적극 권장한다는 여성의 비율은 61%로 남성(83%)보다 낮았다.[27] 소속 회사가 생성형AI 활용법에 대한 직원 교육에 투자한다는 여성의 비율 또한 49%로 남성(79%)보다 낮았다. 이러한 간극은 그저 인식의 문제일 수도, 직장 내 훈련 프로그램이나 정책에 실제로 성별 격차가 있기 때문일 수도 있지만, 기업들이 주의를 기울여 개선 노력을 펼쳐야 하는 사안임은 분명하다.

남성 편중의 AI 부문 인력 구조, AI 모델의 편향성 심화 리스크

하지만 테크 산업 내에서는 이러한 성별 격차가 거의 나타나지 않으므로, 테크 부문에 종사하는 여성이 생성형AI에 대한 여성 전반의 인식을 바꾸는 단서가 될 수 있다. 당연히 AI 산업에 종사하는 근로자들은 생성형AI를 활용하는 비율이 상당히 높다. 딜로이트 서베이에 따르면, 테크 산업에 종사하는 여성 70%, 남성 78%가 생성형AI를 실험적으로 또는 업무용으로 사용하고 있다고 답했다. 이는 비테크 부문의 여성 32%, 남성 40%보다 훨씬 높은 수준이다.[28] 특히 테크 부문 여성은 생성형AI를 실험적으로 사용하는 데서 업무용으로 사용하는 단계로 남성보다 훨씬 빠르게 넘어가고 있다는 점이 눈에 띈다. 서베이 결과 업무용으로 사용하는 여성의 비율이 44%로 33%를 기록한 남성보다 높았다. 테크 부문에서는 남녀 모두 생성형AI로 얻을 수 있는 이점에 대한 기대가 높게 나타났다. 테크 부문 종사자 10명 중 7명은 생성형AI로 인해 1년 후 생산성이 크게 향상될 것이라고 답했다.[29]

또한 테크 부문에서는 생성형AI에 대한 신뢰에 성별 격차가 나타나지 않았고, 남녀 모두 비테크 부문 사용자보다 높은 신뢰를 보였다. 생성형AI 기업들이 자신의 데이터를 안전하게 보호할 것이라 '매우 신뢰한다' 또는 '극히 신뢰한다'고 답한 비율이 남녀 모두 40%를 넘었다.[30] 또한 온라인 서비스로 얻는 이점이 개인정보보호 침해 우려를 상쇄한다는 비율이 남녀 모두 75%로, 비테크 부문 여성 54%, 남성 60%를 훨씬 웃돌았다.[31]

테크 부문 여성은 비테크 부문 근로자들에 비해 생성형AI의 메커니즘을 더욱 잘 이해하고 이미 업무용으로 활용하는 경우가 훨씬 많기 때문에, 불안 수준이 상대적으로 낮고 생성형AI로 얻을 수 있는 이점에 더욱 초점을 맞추기 때문으로 풀이된다. 또한 테크 부문 여성은 대부분 소속 회사가 생성형AI 활용을 장려(84%)하거나 교육을 제공(72%)한다고 답해, 비테크 부문 여성(55% 및 45%)보다 높은 비율을 보였다.[32]

테크 부문 여성은 이처럼 생성형AI 활용과 인식이 상대적으로 매우 긍정적이지만, AI 산업 종사자의 성별 격차는 여전하다. AI 인력 중 여성의 비율은 약 30%에 그쳐, STEM[과학science, 기술technology, 공학engineering, 수학mathematics] 부문 여성 비율과 비슷한 수준을 보이고 있다.[33] AI 부문 인력의 이 같은 성별 격차는 다양한 영역과 부문에서 AI 시스템을 개발 및 배치할 때 심각한 결과를 초래할 수 있다.

AI 인력의 성별 격차로 발생할 수 있는 가장 시급한 문제는 AI 모델의 성별 편향성이 영구화될 수 있다는 리스크다.[34] 범산업적으로 AI 시스템의 최대 44%가 성별 편향성을 보이고 있어, 실제로 여성을 한층 비주류화, 주변화하는 결과를 낳을 수 있다.[35] 예를 들어, AI 모델의 성별 편향성은 채용 방식에 편향성을 초래하거나,[36] 여성의 헬스케어 품질을 저하시키거나,[37] 여성의 금융서비스 이용을 제한할 수 있다.[38] 또한 딜로이트 연구에 따르면, AI 모델의 편향성은 직원과 고객의 신뢰에 부정적 영향을 주는 것으로 나타났다.[39] 양성 평등을 달성하고 AI로 사회 전체를 이롭게 하려면 AI 산업에 더 많은 여성 인력을 확보하는 것이 중요하다.[40]

결론

테크 산업은 다음과 같은 이유로 여성의 생성형AI 활용과 인식을 개선하는 데 많은 노력을 기울여야 한다. 첫째, 여성이 소비지출의 대부분을 결정하거나 영향력을 행사하므로, 여성이 생성형AI를 일상적으로 사용하지 않는다면, 향후 출시될 AI 제품 및 서비스가 매출 목표를 달성하지 못할 수 있다. 둘째, 여성 직원이 남성만큼 적극적으로 생성형AI 툴을 사용하지 않는다면, 기업들은 생성형AI에 투자한 만큼의 생산성 개선을 이루지 못할 수 있다. 셋째, 생성형AI는 사용자와의 대화를 통해 얻은 데이터를 수집하고 추론하기 때문에, 여성과의 대화가 상대적으로 적으면 AI 모델의 편향성이 심화될 수 있다.[41] 마지막으로, 새로운 생성형AI 활용사례에 여성의 개입이 충분하지 않으면, 여성은 헬스케어 챗봇 상담과 같이 기술이 가져올 이점을 충분히 누릴 수 없고, 기존의 불평등이 심화될 것이다.[42]

생성형AI에 대한 여성의 신뢰를 강화하려면, 테크 기업들이 나서서 잠재적 리스크를 완화해야 한다. 딜로이트 서베이에 따르면, 테크 기업들의 개인정보보호 및 데이터 보안 정책의 투명성 개선과 소비자들의 개인정보보호 관리 용이성 개선이 소비자 신뢰에 영향을 주는 것으로 나타

났다.[43] 따라서 테크 기업들은 엄격한 데이터 보안 조치를 우선시하고 데이터 관리 방식을 더욱 효율적으로 설명해야 한다. 어떤 데이터가 어떻게 수집되는지 이해하기 쉬운 방식으로 소비자들에게 설명하고, 소비자들이 필요한 정보에 기반해 개인정보를 손쉽게 관리할 수 있도록 서비스를 개선하면, 신뢰를 구축할 뿐 아니라 경쟁우위도 점할 수 있다. 생성형AI의 리스크를 완화하기 위해 정부의 역할도 중요하다. 딜로이트 서베이 결과, 기업들이 소비자 정보를 수집하고 활용하는 방식에 대한 규제가 강화돼야 한다는 응답자가 84%에 달했다.[44]

각 산업의 기업들은 남녀 직원 모두 생성형AI를 충분히 활용하도록 적극 장려해야 이에 따른 이점을 실현할 수 있다. 문서 편집, 검색, 자료 요약, 조사 등 일반적인 업무용 활용사례 외에도 산업 고유의 활용사례를 적극 수용할 필요가 있다.[45] 직원들의 생성형AI 활용을 극대화하려면 교육 프로그램을 수립하는 것도 중요하다.

소비자들의 생성형AI 활용에 있어서 성별 격차를 좁히는 것만큼 AI 산업 인력의 성별 격차를 줄이는 것도 중요하다. AI 산업 내 여성의 다양성과 포용성을 개선하기 위해, 직원들의 니즈를 반영한 업무환경을 조성할 필요가 있다. 최근 AI 산업에 종사하는 여성에 대한 연구에 따르면, 직장 만족도의 가장 큰 요인으로 유연한 근무시간과 원격근무 등 일과 삶의 균형이 꼽혔다.[46] 또한 여성 인력은 여성 리더십, 임금과 승진의 투명성, 괴롭힘과 남용에 대한 무관용 정책을 갖춘 기업을 선호하는 것으로 나타났다.[47] AI 부문의 여성 인력을 늘리려면, 여성이 AI 스킬과 경쟁력을 갖추도록 교육과 훈련 기회를 확대해야 한다. AI 부문 여성이 경험을 공유하고 상호 지원하는 멘토십 및 네트워킹 프로그램을 늘리고, 펀딩 지원을 통해 AI 연구 및 혁신 프로젝트에 여성의 참여를 늘리는 것도 도움이 된다. 생성형AI 개발에서 여성의 역할이 더욱 중요해지는 만큼, 여성의 더욱 적극적인 관심과 참여를 유도할 수 있는 활용사례와 시스템이 시급하다.

03
생성형AI 스마트폰, 세상을 바꿀 또 한 번의 격변 촉발

> 특화 반도체칩과 광범위한 모바일 운영체제OS 통합으로 스마트폰이 이제 스마트를 넘어 지능화된 기기로 변모하고 있다. 시장의 성장은 소비자들이 이러한 변화를 얼마나 빨리 받아들일 것인가에 달려 있다.

스마트폰은 전 세계에서 가장 많이 사용되는 소비자 기술이다.[1] 스마트폰은 여타 많은 기기들을 통합했고, 반대로 스마트폰의 소형화된 첨단 부품들은 수많은 소비자 및 산업용 기기의 등장을 촉발했다.[2] 손 안에서 모든 것을 처리할 수 있는 편리함과 유용성으로 인해 소비자 행태와 산업의 경쟁 양상이 일변했다. 하지만 최근 스마트폰 혁신은 시장의 흥분을 일으키지 못한 채, 혁신이라기보다 지금까지 등장한 기능의 적재에 불과하다는 평가를 받고 있다.

하지만 이처럼 동면에 들어간 듯했던 스마트폰 생태계가 생성형AIgenerative AI를 경험의 중심으로 끌어오는 차세대 OS와 첨단 반도체칩에 힘입어 스마트폰의 정의를 재정립하고 있다.[3] 생성형AI 기능을 탑재한 스마트폰이 속속 출시되며,[4] 스마트폰 생태계가 다시 한번 요동치고 있다. 하지만 변화와 흥분은 리스크를 수반하게 마련이다.

딜로이트는 2025년 글로벌 스마트폰 출하량이 전년비 7% 증가해, 약 5% 증가했던 2024년에 비해 증가세가 가속화될 것으로 전망한다.[5] 통상적 기기 업그레이드 주기가 겹친 데다, 온디바이스on-device 생성형AI를 지원하는 차세대 스마트폰에 대한 관심이 높은 얼리 어답터 소비자와 개발자들의 업그레이드 수요가 증가할 것으로 전망된다. 2025년 말에 이르면 전체 스마트폰 출하량에서 생성형AI 지원 스마트폰이 30% 이상을 차지할 것으로 예상된다.[6]

생성형AI를 둘러싼 흥분은 여전히 가시지 않고 있다. 하지만 생성형AI가 기대만큼의 혁신을 가져올지, 소비자들이 스마트폰과 소통하는 새로운 방식을 어떻게 받아들일지는 아직 미지수다.[7]

프리미엄 생성형AI 스마트폰, 업그레이드 주기 앞당길까?

주요 스마트폰 제조사들은 단기적으로 주요 스마트폰 생성형AI 통합을 통해 프리미엄 모델에 대한 수요 촉발을 꾀하려 할 것이다. 시장이 포화 지점에 이르면서, 스마트폰 판매는 2022~2023년 2년 연속 감소했다.[8] 현재 전 세계 인구의 절반을 넘는 약 50억 명이 스마트폰을 소유하고 있는 것으로 추산된다.[9] 게다가 최근 수년간 모델 업그레이드 주기가 길어지고 있다. 평균 업그레이드 주기는 2~3년이지만, 최근 인플레이션 압력이 높아져 소비자들의 재량지출이 제약을 받고 있다.[10] 이와 동시에 몇 년간 사

용할 제품이므로 기왕이면 상위 모델을 선택하는 추세가 강해지고 있다.[11] 따라서 스마트폰 업계는 단순히 하드웨어 업그레이드뿐 아니라 경쟁력 있는 가치와 편리성으로 무장한 사용자 경험을 제공해야 한다.

2024년 1분기에는 스마트폰 판매가 크게 증가했다. 소비자 신뢰도가 회복된 데다 프리미엄 생성형AI 스마트폰에 대한 관심이 막 싹트기 시작한 덕분이다.[12] 딜로이트 2024년 서베이에 따르면, 커넥티드 기기 구매 시 경제적 요인의 영향을 받는다는 소비자가 줄었다.[13] 특히 유럽은 2024년 2분기까지 스마트폰 판매가 꾸준한 증가세를 보여 이러한 서베이 결과를 뒷받침했다.[14] 따라서 2025년에는 스마트폰을 교체하는 소비자가 늘고, 이들 중 상당수는 생성형AI 기능을 탑재한 고가의 프리미엄 스마트폰을 택할 가능성이 크다.

생성형AI가 스마트폰 업그레이드의 동인이 될 수 있지만, 소비자들의 양상은 시장과 연령대에 따라 다르게 나타날 수 있다. 딜로이트 서베이에 따르면, 미국 소비자 중 생성형AI 기능을 써보고 싶어서 계획보다 빨리 스마트폰을 교체한다는 비율이 7%에 그쳤으나, 연령대를 스마트폰 의존도와 신기술 수용도가 높은 24~45세로 좁히면 그 비율은 50%로 뛰었다.[15] 하지만 딜로이트 영국 소비자 대상 서베이에서는 응답자 4%만이 생성형AI를 매일 사용한다고 답했고, 생성형AI가 유용하다는 응답자는 23% 그쳤으며 19%는 생성형AI의 답변에 만족하지 못한다고 답했다.[16]

생성형AI가 스마트폰 업그레이드 주기를 앞당길지는 생성형AI가 제시하는 가치와 유용성에 달려 있다. 2025년 스마트폰이 생성형AI의 유용성을 시험에 들게 할 매개체가 될 수 있다.

PC의 생성형AI 기능

스마트폰과 마찬가지로 생성형AI 특화 온디바이스 칩을 탑재한 차세대 PC 시장 또한 사용자 경험, 유용성, 가치, 하이퍼스케일 생성형AI의 진화를 형성하는 전반적인 압력 등이 다각도로 작용한다.

딜로이트 서베이에 따르면, 소비자들은 생성형AI 지원 PC 구매에 관심을 보였다. 미국 응답자 중 34%가 생성형AI 지원 PC로 교체하기 위해 계획보다 일찍 노트북을 교체할 의향이 있다고 답했다. 딜로이트 추산에 따르면, 연간 PC 판매량의 약 50%가 개인 소비자들에게 판매되므로, 소비자들의 이러한 행태는 매우 중요한 고려 요인이다.[17] 기업 구매자들은 어떤 생성형AI 코프로세서 PC 모델을 구매해야 사업적 타당성이 가장 높을지 아직 불확실하다는 입장이다. 관련 시장에서 다양한 PC OEM사들이 다양한 옵션과 가격을 제시하고 있기 때문이다.[18]

향후 최첨단 PC에는 특수 실리콘칩으로 생성형AI 기능이 탑재될 것이다. 최근 연구에 따르면, 2028년에 이르면 판매되는 모든 PC 중 80%에 이러한 실리콘칩이 탑재될 것으로 전망됐다.[19] 또 다른 연구에 따르면, 2024년 2분기에 AI 기능 탑재 PC 약 900만 대가 출하된 것으로 추산됐다. 하지만 이 중 생성형AI 작업부하를 운영할 만큼 강력한 신경망 처리장치NPU가 탑재된 PC가 얼마나 되는지는 불확실하다.[20] 결론적으로 잠재적 구매자들은 더욱 성능이 뛰어난 차세대 생성형AI PC의 출시를 1년여간 기다린 후 PC 교체에 나설 가능성이 크다.

딜로이트는 2024년에 판매된 PC 중 약 30%에 온디바이스 생성형AI 프로세싱 기능이 탑재돼 있을 것으로 추산하며,[21] 2025년에는 그 비율이 50%에 육박할 것으로 전망한다.

2024년 기준 판매량이 2억6,100만 대로[22] 12억3,000만 개[23]를 기록한 스마트폰에 비하면 규모가 작지만, 컴퓨터는 평균 판매가가 더 높기 때문에 판매액은 판매량보다 격차가 적다. 2024년 컴퓨터 판매액은 약 2,200억 달러, 스마트폰 판매액은 5,200억 달러로 추산된다.[24]

생성형AI PC 출시가 컴퓨터 시장에 어떠한 영향을 미칠지는 아직 알 수 없다. 다만 평균 판매가가 약 15% 상승할 것으로 예상된다.[25] 하지만 PC 판매는 한 자릿수 증가율에 그칠 것으로 전망된다.[26]

스마트폰과 PC 부품이 발전하면 공급망이 형성돼 가격이 내려가고, 이러한 공급망 속에서 동일 부품이 여타 다수의 기기 시장으로도 유입된다. 생성형AI 관련 부품 또한 이 같은 경로를 밟아 여타 커넥티드 기기에서도 일반화될 것으로 기대된다.

생성형AI로 지능화된 스마트폰

스마트폰의 '스마트'라는 표현은 앱을 구동할 수 있는 커넥티드 기기라는 의미로 통용된다. 하지만 생성형AI가 탑재되면 스마트폰은 더욱 개인화되고 사용자의 의도와 상호작용을 인식하고 대화형 인터페이스로 더욱 친밀해질 수 있다. 음성 비서 등 이전의 개인화 노력은 기대에 미치지 못했지만, 일부 사용자들은 이미 최신 대화형 대규모언어모델LLM과 '관계'를 쌓고 있다.[27] 이러한 변화는 디지털 시스템과의 인터페이스 방식으로서 대화형 AI와 새로운 상호관계의 패러다임이 될 수 있다. 이와 동시에 사용자를 대신해 행동하는 법을 학습한 신뢰할 수 있는 지능형 에이전트의 새로운 모델이 될 수 있다.

온디바이스 생성형AI 모델은 "오후 2시 약속에 늦지 않으려면 몇 시에 출발해야 하지?"라는 질문에 답할 수 있다. 사용자의 의도를 추론하고 사용자의 캘린더, 위치, 제한 시간 내 목적지까지 이동하는 최상의 경로 등을 이해할 수 있기 때문이다. 온디바이스 모델은 초당 30 테라 연산TOPS, tera operations per second, 1초에 1조 번의 연산 수행을 넘는 성능[28]을 갖춘 NPU를 활용해 온디바이스 추론을 지원하기 때문에, 범위가 좁은 태스크를 수행하는 데 탁월하다. 만약 특정 질문이 로컬의 영역을 넘어선다고 판단하면 온디바이스 모델이 해당 태스크를 성능이 더욱 큰 클라우드 기반의 모델로 전송해 답을 얻는다. 고성능 모바일 컴퓨팅을 기반으로 클라우드의 대규모 모델에 직접 접속하는 이러한 하이브리드 방식을 활용하면, 스마트폰만으로도 더욱 즉각적이고 안전한 상호작용을 할 수 있다.[29]

소규모 온디바이스 생성형AI 모델을 이용하면 필요 시 사용자 상호작용과 데이터를 로컬에 안전하게 저장하고, 실시간 번역 등 매우 빠른 답변이 필요할 때 지연속도를 한층 줄일 수 있다.[30] 이러한 경험은 사용자의 신뢰와 유용성을 끌어올리는 데 도움이 된다. 또 스마트폰 기업 입장에서는 사용자 상호작용에서 수집한 데이터를 기반으로 플라이휠 효과flywheel effect를 얻어 사용자가 더욱 만족할 만한 답변을 하도록 로컬 및 클라우드 모델을 개선할 수 있으며 더욱 심도 깊은 비즈니스 인사이트도 얻을 수 있다.

더욱 장기적으로는 소비자 상호작용의 명실상부한 중심점이 된 스마트폰이 맞춤화와 지능화를 한층 강화해 개인의 행태에 대한 이해와 개인

의 니즈에 대한 예측을 더욱 심화할 수 있다. 스마트폰이 이처럼 '에이전틱'agentic 기기로 진화하면 단순한 '스마트' 기기에서 '지능형' 기기로 완전히 변모할 것이다.

2025년 한 해는 소비자들이 초기 생성형AI 기능의 가치를 시험하고 파악하며 새로운 경험을 얼마나 빨리 받아들이는지를 파악하는 전환기가 될 것이다. 스마트폰 기업들은 향후 수 개월 생성형AI 기능을 추가한 모델을 출시하겠지만, 생성형AI를 전면에 내세운 모델의 출시에는 좀 더 시간이 걸릴 것이다.[31] 2025년은 클라우드와의 하이브리드 방식보다는 소규모 온디바이스 모델의 성능과 한계를 시험하는 한 해가 될 것으로 보인다. 이렇게 되면 생성형AI를 둘러싼 경제학이 변화할 수 있다. 비용이 많이 드는 데이터센터 대신 온디바이스로 수행되는 생성형AI 태스크가 늘어나면, 생성형AI의 자본 집약적 부담이 완화될 수 있다.

스마트폰 산업, 생성형AI에 대한 열기가 사라져도 투자를 지속할 수 있는가?

시장에서는 프런티어 모델frontier model, 최첨단 고성능 AI 모델이 뛰어난 제품 적합성product fit*을 갖춰 높은 비용이 제 값을 해야 한다는 강력한 압력이 작용하는 한편, 기업들은 모델 개발과 운영을 지속하기 위해 비용 효율성을 더욱 강화해야 하는 압력을 받고 있다.[32] 주요 스마트폰 기업들은 현재의 프런티어 모델을 개발하기 위해 수십 억 달러를 투자했고, 대규모 수요를 충족하는 데 필요한 데이터센터 구축에 또 수십 억 달러를 투자하고 있다.[33] 일부 연구에 따르면, 생성형AI 지원을 위해 매년 6,000억 달러가 투자되는 것으로 추산됐다.[34] 하지만 이러한 자본 집약적 투자는 경제적 가치를 얻지 못한다면 지속할 수 없으므로, 제품 적합성을 개선하려는 움직임이 이어질 것이다.

* 제품 적합성(product fit)은 특정 제품이 사용자에게 실제 가치를 제공하는지, 시장에서 수요와 일치하는지를 측정하는 데 사용되는 개념으로, 목표 시장, 고객의 요구, 특정 문제를 해결하는 데 얼마나 잘 부합하는지를 평가하는 데 활용된다.

우선 생성형AI 모델의 규모를 축소해 입력 데이터의 양을 줄이고, 작업부하의 범위에 기반해 데이터를 분산시키면 사용 목적에 따라 추론 태스크의 규모를 조정해 비용을 줄일 수 있다. 또한 소비자와 근로자가 요구하는 태스크 대부분이 생성형AI에 그대로 노출되므로, 더욱 비용이 적게 들고 에너지 효율적인 소규모 모델을 활용하면 이러한 문제도 해결할 수 있다.

하지만 얼마나 많은 추론 태스크를 온디바이스에서 수행하게 될지는 아직 불확실하다. 현재 소비자들의 생성형AI 상호작용과 결과에 대한 기대는 퍼블릭 클라우드 기반 모델의 수행 능력에 맞춰져 있다. 어떤 종류의 태스크와 프롬프트가 로컬에서 무료로 안전하게 수행되고, 어떤 태스크가 네트워크를 통해 클라우드로 이동해야 하는지를 소비자들이 제대로 이해하기까지는 시간이 걸릴 수 있다. 따라서 온디바이스와 클라우드를 혼합한 하이브리드 방식의 대화형 에이전트와 소통한다는 새로운 패러다임을 소비자들이 어떻게 받아들일지는 아직 미지수다.

생성형AI 스마트폰의 범용화를 가로막는 다수의 장애물

딜로이트 서베이에 따르면, 미국 소비자 중 38%가 생성형AI를 사용하고 있으며, 사용자 중 63%는 생성형AI가 기대 이상이라고 답했다.[35] 생성형AI를 이미 사용하는 소비자들은 이미 마법과 같은 경험을 하고 있겠지만, 스마트폰 기업들은 새로운 모델의 높은 가격을 정당화할 수 있도록 더 많은 소비자층에 어필할 수 있는 유용성을 증명해야 하는 과제를 안고 있다.

사용자들은 이전과는 완전히 다른 새로운 상호작용을 원하기 때문에 생성형AI 스마트폰을 어떻게 활용해야 하는지를 둘러싸고 혼란이 발생할 수 있다. 사용자들은 자신의 캘린더를 스스로 관리하려는 AI 비서가 꺼림칙할 수 있다.[36] 이 외 배터리 소비가 빨라질 수도 있고, 퍼블릭 통합 모델 사용료가 추가로 부과될 수도 있으며, 인식하지 못한 오류로 인해 중요한 일을 망칠 수도 있다. 이러한 혼란들이 뒤섞여 사용자와 AI 에이전트, 퍼블릭 모델 사이 신뢰를 구축하는 데에는 오랜 시간이 걸릴 것이다. 하지만 이렇게 공들여 구축한 신뢰가 무너지는 것은 한 순간이다.

스마트폰 기업들은 차세대 프런티어 모델로 더욱 큰 가치가 창출될 것이라 기대하지만, 프런티어 모델의 성능이 계속 발전될지, 정체 또는 악화될지 아직 알 수 없다. 우선 갈수록 탐욕스러울 정도로 더욱 많은 데이터를 필요로 하는 AI 모델의 재훈련을 지속할 수 있는가?[37] AI 모델이 생성한 합성 데이터로 다시 AI 모델을 훈련시킨다는 솔루션도 제시됐지만, 이는 반복될수록 AI 모델의 추론 품질을 저하시킬 수 있다는 문제가 있다.[38] 또 데이터와 훈련, 추론에 들어가는 비용을 늘리지 않고도 모델의 성능을 개선할 수 있는가? 자본과 데이터 투자가 줄어도 모델의 성능을 개선할 수 있는 방법이 있는가? 인내심을 잃은 투자자들은 AI 기술이 이러한 문제를 해결할 만큼 발전하기 전에 수익을 창출하라는 압력을 가할 가능성이 크다.

규제당국도 생성형AI의 발전에 영향을 미칠 수 있다. 딥페이크와 잘못된 정보, 인간과 흡사한 봇bot을 이용한 사기 행위 등에 대한 광범위한 안전장치의 마련이 시급하다. 대화형 봇은 사용자와 친밀한 관계를 형성해 사용자의 생각과 이데올로기까지 변하게 할 수 있다.[39] 개인화 대화형 에이전트는 인간 상호작용의 더욱 깊은 영역까지 들어가, 인간에게 긍정적 도움을 줄 수도 있지만 AI와의 소통에 중독되는 사람들이 생길 수도 있다.[40] 또 온디바이스 생성형AI와 외부 모델을 결합하면 보안에 취약한 공격 노출면이 확대될 수 있다.[41] 이에 따라 스마트폰 기업들의 생태계 보안 강화 및 규제당국의 안전장치 강화 노력이 강화될 것으로 예상된다.

결론

또 한 번 시장의 격변을 불러올 '차세대 스마트폰'이 여기저기서 거론되고 있지만 아직 현실화된 것은 아니다. 수십 억 명이 사용하는 스마트폰은 여전히 압도적인 디지털 기기이며 새로운 서비스와 상호작용을 실험

하는 대형 시험대 역할을 한다. 2025년에도 프리미엄 스마트폰과 PC를 통해 생성형AI라는 신기술과 소통하는 소비자의 수가 급증할 것으로 예상된다. 소비자들은 익숙한 기기를 통해 생성형AI를 시험삼아 사용해 보고 그 가치를 확인함과 동시에 한계를 시험할 것이다. 소비자들이 이러한 경험에 만족하면, 스마트폰은 더욱 매력적인 기기와 더욱 확장된 플랫폼으로 진화해 완전히 새로운 유형의 사용 방식과 기회를 창출함으로써 소비자 기기 시장에 새로운 붐을 일으킬 수 있다. 하지만 이러한 여정은 시간이 필요하고, 2025년은 사용자들이 새로운 패러다임에 적응하는 시간이 될 것이다.

향후 수 년간 스마트폰 OS는 다양한 상호작용을 통합할 것이다. 원격 링크 대신 로컬 모델이 정보를 요약해 제공하는 차세대 대화형 검색 기능이 탑재되는 등 서비스 제공업체와 정보 출처 간 연결고리가 점차 사라질 수 있다. 개인화된 에이전틱 AI를 사용하는 소비자들이 증가하면, 디지털 상호작용의 성격이 변모해 직접적 인터페이스 대신 사용자의 기기에서 더 많은 태스크를 수행할 수 있다. 또한 다양한 상호작용을 통합하는 스마트폰을 통해 사용자가 의식하지 않아도 주변의 컴퓨팅 환경을 통해 스스로 동작하는 앰비언트 컴퓨팅ambient computing과 디지털 정보와 물리적 공간이 상호작용하는 공간 컴퓨팅spatial computing이 한층 일반화될 수 있다.

스마트폰 업계가 이처럼 다시 한번 수요를 촉발하기 위한 혁신을 준비하는 가운데, 대규모 AI 모델 훈련과 운영에 투자된 자본과 에너지 비용을 상쇄해야 한다는 경제적 압박이 만만치 않다. 우선 소규모 AI 모델과 하이브리드 아키텍처부터 시도함과 동시에 어떤 생성형AI 작업부하에 어떤 종류의 계산 오버헤드computational overhead*가 필요한지를 심층적으로 파악해야 한다. 한편 기후 불확실성과 우려를 피해갈 수 없는 지금, 생성형AI 데이터센터 확장으로 이미 에너지와 물 사용량이 늘어 소비자와 정부가 부담해야 할 에너지 비용이 늘고 있다.[42] 생성형AI의 가치가 경제적 부채를 극복할 정도로 상승했다 하더라도, 아직 에너지 부채는 갚지 못하는 실정이다.

* 계산 오버헤드(computational overhead)는 AI 알고리즘이 실제 태스크를 수행할 때 추가적으로 소모하는 시간과 메모리, 처리 능력 등 계산 자원을 의미하며, 작업 효율에 영향을 줄 수 있다.

2024년 말 현재 하이퍼스케일러hyperscaler, 대규모 데이터센터 운용 업체와 스마트폰 생태계 플레이어들, 신규 퍼블릭 모델 기업들이 제공하는 이점이 경제 전반에 가치를 창출할 것이라는 전망이 우세하다. 하지만 정작 투자를 단행한 당사자들은 얼마나 많은 가치를 확보할 수 있는가? 과거 통신사들과 초기 인터넷 회사들은 막대한 자본지출을 들여 인프라를 구축해, 차세대 혁신 기업들에 동력을 제공하는 역할을 했다. 생성형AI 하이퍼스케일러들도 같은 길을 걸을 것인가?[43]

생성형AI의 등장부터 배치, 범용화를 주도하는 것은 인터넷 범용화 이후 인류의 가장 위대한 실험에 속한다. 최종 도착지는 아직 불확실하지만, 생성형AI라는 혁신으로 수많은 신기술과 행태, 사업모델이 창출될 수 있는 여정이 시작됐다.

04
자동화의 새로운 시대 여는 에이전틱 AI

> 자동화 생성형AIgenerative AI 에이전트, 즉 에이전틱 AIagentic AI가 지식 근로자의 생산성을 향상하고 온갖 유형의 업무 흐름 효율성을 개선할 수 있다. 하지만 완전한 자동화가 범용화되기까지는 아직 갈 길이 멀다.

에이전틱 AI는 인간의 감독을 거의 또는 전혀 받지 않고도 복잡한 태스크를 완수하고 목표를 달성할 수 있는 소프트웨어 솔루션으로, 흔히 '에이전트'agent로 불리는 현재의 챗봇chatbot이나 코파일럿co-pilot과는 다르다. 에이전틱 AI는 지식 근로자들의 생산성을 끌어올리고 사업 전반의 다단계 프로세스를 자동화할 수 있다. 딜로이트는 2025년 생성형AI를 도입한 기업들 중 25%가 에이전틱 AI를 시범 도입하거나 기술 검증에 나설 것으로 전망한다. 또한 2027년에 이르면 그 비율은 50%로 늘어날 것으로 예상한다.[1] 일부 산업이나 활용 사례의 경우 2025년부터 특정 에이전틱 AI 앱이 실제로 업무에 적용될 수 있으며, 이러한 추세는 하반기로 갈수록 뚜렷해질 것으로 전망된다.

스타트업과 기존 테크 기업들이 수익 잠재력을 기대하고 에이전틱 AI 개발에 힘을 쏟으며 이러한 추세를 뒷받침하고 있다. 지난 2년간 에이전틱 AI 개발 스타트업에 유입된 투자자본은 20억 달러가 넘는다. 투자는 기업용 시장에 주력하는 스타트업에 집중됐다.[2] 한편 다수의 테크 기업들과 클라우드 기업들이 자체 에이전틱 AI 서비스를 개발하고 있다. 이들은 또한 전략적 인수에 나서거나, 전면 인수 대신 스타트업들로부터 에이전틱 AI 라이선스를 구매하거나 인력을 확보하는 경우도 늘어나고 있다.[3]

인간의 개입 없이 스스로 태스크를 완수하는 에이전틱 AI

생성형AI 챗봇과 코파일럿은 첨단 기술이다. 인간과 직관적으로 소통하고 복잡한 정보를 가공하고 콘텐츠를 생성한다. 하지만 에이전틱 AI와 달리 자율성은 없다. 챗봇과 에이전틱 AI는 대규모언어모델LLM이라는 동일한 기초 모델을 공유하지만, 에이전틱 AI는 추가 기술이 적용돼 인간의 감독이 거의 또는 전혀 없이도 독립적으로 행동하고 태스크를 단계별로 수행해 완수한다. 에이전틱 AI는 소통하는 데 그치지 않고, 사용자 대신 사용자보다 훨씬 효율적으로 추론하고 행동한다.

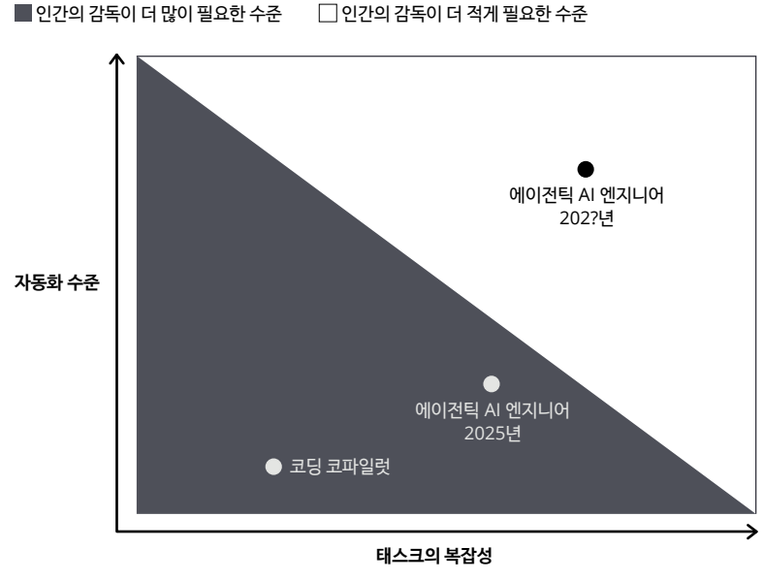
이름에서 알 수 있듯이 에이전틱 AI는 '에이전시'agency로서 행동하고 어떤 행동을 할지 스스로 결정한다.[4] 에이전시는 독립적으로 행동하고 결정할 수 있는 자율성을 의미한다.[5] 즉 목표를 달성하기 위해 스스로 계획하고 실행할 수 있다는 의미다.[6] 목표는 인간이 설정하지만, 그 과정은 에이전틱 AI가 정한다.

에이전틱 AI와 코파일럿 및 챗봇의 차이는 다음과 같이 설명할 수 있다. 코드를 테스트하고 제시함으로써 소프트웨어 개발자들을 돕는 코파일럿은 지금까지 가장 성공적인 생성형AI 활용 사례로 꼽힌다.[7] 경험 많은 소프트웨어 엔지니어들의 생산성을 끌어올리고 주니어 엔지니어들의 효율성을 개선한다. 코파일럿은 다양한 언어의 자연어 프롬프트를 코드 제안으로 전환해 코드의 일관성을 테스트한다. 하지만 엔지니어들의 프롬프트에 응답할 뿐 자율성을 가지고 있지는 않다. 하지만 에이전틱 AI와 일하는 엔지니어는 한 단계 더 나아갈 수 있다. 사람 엔지니어가 프롬프트로 소프트웨어 개발 아이디어를 입력하면, 에이전틱 AI 엔지니어가 이러한 아이디어를 실행 가능한 코드로 전환한다. 소프트웨어 개발 과정에서 다수의 단계가 자동화되는 것이다.

실제 사례로, 2024년 3월 코그니션 소프트웨어Cognition Software사가 수 천 가지의 결정을 내려야 하는 복잡한 엔지니어링 태스크를 추론, 계획, 완수할 수 있는 자동화 소프트웨어 엔지니어 '데빈'Devin을 출시했다.[8] 데빈은 사람 프로그래머의 자연어 프롬프트에 기반해 인간의 개입 없이 앱 개발, 코드베이스 테스트 및 수정, LLM 훈련 및 재훈련 등 프로그래밍을 수행하도록 설계됐다.[9] 2024년 여름에는 데빈의 오픈소스 버전 기업용 코드 도우미라 할 수 있는 코디움Codeium도 출시됐다.[10]

하지만 에이전틱 AI 엔지니어들도 코파일럿과 비슷한 취약점을 가지고 있다.[11] 현재 기술로는 인간의 감독 없이 완전히 자율적으로는 고사하고 일부 자율적으로라도 태스크를 맡기기에 오류가 지나치게 많다. 최근 벤치마킹 테스트에서 데빈은 실제 코드 저장소에서 가져온 깃허브GitHub 이슈 중 약 14%를 해결해, LLM 기반 챗봇보다 두 배 뛰어난 실력을 보였으나,[12] 여전히 완전 자동화라고 하기에는 턱없이 부족하다. 하지만 빅테크 기업들[13] 과 스타트업들이 에이전틱 AI 엔지니어들의 자율성과 신뢰성을 개선하기 위해 주력하고 있는 만큼, 에이전틱 AI는 사람 엔지니어 워크로드의 일부를 처리할 수 있는 수준에 이르렀다(그림 1).

그림 1
에이전틱 AI의 발전



출처: 딜로이트 분석

인간의 생산성 증강하는 에이전틱 AI

에이전틱 AI 소프트웨어 엔지니어는 자율성을 지닌 생성형AI가 일의 방식을 어떻게 바꾸는지를 보여주는 한 가지 예시일 뿐이다. 에이전틱 AI가 한층 발전하면 커다란 변화를 가져올 것이다. 지식 근로자 수가 미국에서만 1억 명이 넘고, 전 세계적으로는 12억5,000만 명이 넘는다.[14] 미국에서 지식 노동의 생산성을 측정하는 요소 생산성factor productivity[15]은 1987년부터 2023년까지 0.8%, 2019년부터 2023년까지 0.5% 증가하는 데 그쳐 거의 정체 양상을 보였다.[16] 경제협력개발기구OECD 회원국들도 마찬가지다.[17] 자동화를 통해 지식 노동의 생산성을 개선하려는 시도는 일부의 성공만을 거뒀다. 그리고 수많은 기업들은 여전히 지식 근로자가 부족한 실정이다. 고객 서비스 인력부터 반도체 엔지니어까지 거의 모든 지식 근로자가 부족해, 신규 인력이 투입되면 즉각 생산성을 도출해야 한다.

자동화를 위한 전문가 시스템expert system과 로보틱처리자동화RPA는 단계가 모호해지거나 여러 단계를 거쳐야 하는 경우 불안정해질 수 있다. 기존 머신러닝에 기반한 시스템은 특정 용도에 맞춰 광범위한 훈련이 필요하다. 하지만 LLM 기반의 에이전틱 AI는 머신러닝이나 딥러닝보다 유연하게 다양한 활용 사례에 적용할 수 있다.

에이전틱 AI는 LLM의 역량을 대폭 끌어올려 기업들의 생성형AI에 대한 투자 가치를 증명할 수 있다. 기업 리더들은 생성형AI 툴을 도입해 즉각적 효과를 기대했지만, 그에 따른 비즈니스 가치를 수량화하기가 쉽지 않다.[18] 데이터 기반, 리스크 및 거버넌스 정책, 인력 격차 등으로 인해 생성형AI 이니셔티브의 규모를 확대하기도 쉽지 않다. 생성형AI 시범 도입 사례 중 본격적 도입으로 이어진 것은 30%에 불과하다.[19] 게다가 생성형AI가 도출한 결과물에 대한 신뢰 부족과 생성형AI의 실수가 실제 세계에 가할 수 있는 피해 등도 기업의 전면적 도입을 주저하게 만드는 요인으로 작용한다.[20]

에이전틱 AI를 개발하고 실행하는 기업들은 생성형AI에 수반되는 이 같은 문제뿐 아니라 추론, 행동, 협업, 창작이 가능한 봇을 개발하는 과정의 복잡성도 유념해야 한다. 가장 중요한 것은 생성형AI 에이전트의 신뢰성이 보장돼야 한다는 점이다. 대부분의 경우 태스크를 제대로 완수하는 것만으로는 충분하지 않다. 2024년 말에는 생성형AI를 완전히 신뢰할 수 있다는 일련의 고무적인 활용 사례들이 나타나 2025년 초 전면 도입의 가능성이 커지고 있다.

신뢰성이 개선된 에이전틱 AI의 잠재적 이점을 생각하면 이러한 노력은 충분히 기울일 만한 가치가 있고, 초기 성과도 고무적이다. 기업들은 생성형AI 모델을 여타 AI 기술 및 훈련 기법과 접목해 LLM의 성능을 강화시키는 방법을 모색하고 있다. 완전히 자율화되고 신뢰할 수 있는 에이전트가 최종 목표이기는 하지만, 정확성과 독립성을 점증적으로 개선하는 과정에서 전반적으로 생성형AI의 생산성과 효율성을 개선하는 초

기 목표를 달성할 수 있다.[21] 에이전틱 AI는 범산업적 솔루션으로도 산업 특화 솔루션으로도 광범위하게 활용할 수 있기 때문에, 당초 기업 리더들이 생성형AI에 기대하던 효과를 기대해볼 만하다.

에이전틱 AI의 메커니즘

에이전틱 AI는 복잡한 태스크를 일련의 단계로 정리해 실행하고, 예기치 않은 장애물도 해결할 수 있다. 또한 활용 사례에 따라 가상 환경이나 물리적 환경, 또는 두 가지 환경을 모두 감지할 수도 있다. 에이전틱 AI는 태스크를 완수하기 위해 어떤 행동을 취해야 할지 정하고, 다른 툴과 데이터베이스, 에이전트를 활용하며, 인간이 정한 목표에 기반해 결과를 제시한다.

에이전틱 AI는 계속 진화하는 신흥 기술이지만 다음과 같이 몇 가지 공통적 특징이 있다.

- **기초 모델에 기반:** 에이전틱 AI는 LLM과 같은 기초 모델에 기반해 추론과 분석을 하고 복잡하고 예측하지 못한 업무 흐름에 적응한다. 이 때문에 에이전틱 AI는 RPA나 전문가 시스템보다 유연하다. LLM의 최신 혁신 중 하나는 증강된 추론 능력과 태스크를 일련의 단계로 쪼개는 능력이다.[22] 하지만 기초 모델 스스로는 여타 환경과 상호작용하고 결정을 내리고 태스크를 실행할 수 없으며,[23] 여타 기술과 역량으로 증강돼야 한다.

- **자율적 행동:** 자율성의 수준은 다양하지만 에이전틱 AI는 스스로 복잡한 태스크를 계획하고 실행하도록 훈련할 수 있다. 추론 토큰reasoning token*을 도입함으로써 사고의 흐름 모델chain of thought model** 이 이전 LLM보다 복잡한 문제를 해결할 수 있다. 사고의 흐름 모델은 응답 속도가 더 느리지만, 문제 해결을 위한 추론 방식이 더욱 정교하고 오류를 스스로 수정할 수 있으며 결론을 도출한 과정을 설명할 수 있다.[24]

  * 추론 토큰(reasoning token)은 AI 모델이 특정 문제를 해결하거나 답변을 생성하기 위해 논리적 추론을 수행할 때 생성 또는 활용되는 텍스트 단위를 뜻한다.
  ** 사고의 흐름 모델(chain-of-thought model)은 AI와 자연어 처리(NLP)가 문제를 해결하기 위해 단계적으로 추론하는 기술을 뜻한다.

- **환경 감지:** 에이전틱 AI는 환경을 감지해 정보를 처리하고 주어진 태스크의 맥락을 이해할 수 있다.[25] 또 첨단 에이전틱 AI는 비디오와 이미지, 오디오, 텍스트, 숫자 등 멀티모달 데이터를 처리할 수 있다.

- **툴 사용:** 에이전틱 AI는 소프트웨어와 기업용 앱, 인터넷 등 여타 툴과 시스템과 상호작용하며 태스크를 완수한다.

- **조율:** 에이전틱 AI는 여타 시스템과 봇에 참여를 지시해 태스크를 완수한다. 이 같은 멀티에이전트 시스템으로 여타 자동화 생성형 AI 에이전트와 협업할 수 있다.

- **메모리 접근:** LLM은 '상태 또는 기억이 없다'stateless고 알려져 있다. LLM은 각각의 상호작용을 독립적으로 처리하며, 상호작용이 완료되면 정보를 재훈련하지 않는다. 하지만 에이전틱 AI는 복구 가능한 메커니즘과 데이터베이스를 갖추고 있어 특정 태스크를 수행할 때 단기 메모리에 접근해 컨텍스트를 유지할 수 있고 장기 메모리에 접근해 과거의 경험으로부터 학습하고 개선할 수 있다.[26]

일부 최신 에이전틱 AI 모델은 사고의 흐름 모델이 도입돼 이전 대규모 모델보다 속도는 느리지만 추론 방식이 더욱 정교해 복잡한 문제에 대해 더욱 고차원적 추론이 가능하다.[27] 또 멀티모달 데이터 분석이 가능해 해석 및 생성할 수 있는 데이터 종류가 확대되기 때문에 유연성이 더욱 뛰어나다. 멀티모달 AI는 컴퓨터 비전(이미지 인식)이나 전사 및 번역과 같은 여타 AI 기술과 결합할 때 더욱 강력한 성능을 발휘한다.[28] 에이전틱 AI와 마찬가지로 멀티모달 AI도 계속 진화하고 있다.

태스크가 자동화 에이전트의 네트워크를 오가며 수행되는 진정한 멀티에이전트 시스템이 한창 개발 중이며, 일부 파일럿 프로젝트가 2024년 말 론칭됐다.[29] 멀티에이전트 모델은 복잡한 환경에서 태스크를 분산하기 때문에 싱글에이전트 모델보다 성능이 뛰어나다.[30] 스타트업과 빅테크 기업들은 기업들이 자체 맞춤형 에이전트를 구축할 수 있는 툴이 포함된 멀티에이전트 생성형AI 시스템을 개발하고 있다.[31]

에이전틱 AI의 유망 활용 사례

빅테크 기업들과 스타트업들은 소프트웨어 개발, 세일즈, 마케팅, 규제 컴플라이언스 등 기능을 일부 자동화하기 위한 초기 단계 에이전틱 AI 솔루션을 개발 중이다. 아직 기술 검증 단계이거나 기업용 도입 준비가 되지 않은 사례도 있지만, 에이전틱 AI는 다음과 같은 활용 사례가 유망하다. 대부분 범산업적 활용 사례이지만, 산업 특화 활용 사례도 제시되고 있다.

- **고객 지원:** 고객 서비스는 필수적이지만 스트레스가 과도한 업무이기도 해, 연간 이직률이 38%에 달한다.[32] 고객 지원 업무의 일부를 효과적으로 자동화하면 인력은 스트레스가 줄고 기업은 더 많은 고객에게 서비스를 제공할 수 있다.[33] 에이전틱 AI는 현재의 고객 지원 챗봇보다 더욱 복잡한 고객 요청에 대응할 수 있으며, 자율적으로 문제를 해결할 수 있다. 일례로, 한 오디오 기업이 에이전틱 AI를 도입해 고객들의 새로운 오디오 장비 설치를 지원하도록 함으로써 통상 인간의 개입이 필요한 몇 가지 단계의 프로세스를 자

동화했다. 만약 인간의 개입이 필요하다면 에이전틱 AI가 관련 정보를 취합하고 상황을 요약해 해당 고객을 사람 직원에게 인계한다.[34] 다음 단계의 고객 지원 에이전틱 AI는 텍스트 기반 챗에 음성과 비디오 등 멀티모달 데이터를 통합한 형태가 될 수 있다.

- **사이버 보안:** 사이버 보안은 전문가 부족난이 심각한 대표적 분야로, 현재 전 세계적으로 400만 명의 인력이 부족한 실정이다.[35] 그러는 동안 생성형AI를 활용해 사이버 보안 시스템을 침투하는 해킹이 기승을 부리고 있다. 에이전틱 사이버 보안 시스템이 개발되면 자동화를 통해 사람 전문가의 효율성이 한층 개선된다. 에이전틱 AI가 자율적으로 공격을 감지해 보고하면, 시스템 보안을 개선하고 사람 전문가의 워크로드를 최대 90% 줄일 수 있다.[36] 또한 소프트웨어 개발 팀을 도와 새로운 코드의 취약성을 파악해, 테스트를 수행하고 개발자와 직접 소통하면서 문제를 해결할 수 있다. 현재 사람 엔지니어들이 직접 수행하는 업무를 자동화할 수 있는 것이다.[37]

- **규제 컴플라이언스:** 각종 산업, 특히 금융서비스와 의료 부문의 기업들은 주기적으로 규제 컴플라이언스 리뷰를 수행해야 한다. 관련 규제의 규모와 복잡성은 증대하는데 컴플라이언스 전문가 부족해 컴플라이언스가 기업들의 심각한 해결과제가 되고 있다. 이에 스타트업들이 규제를 분석하고 문서를 생산하며 기업의 컴플라이언스를 신속하게 평가할 수 있는 에이전틱 AI를 개발하고 있다. 컴플라이언스 에이전틱 AI는 특정 규제를 인용하고 선제적 분석을 제시하며 사람 전문가들에게 조언을 제공할 수 있다.[38] 현재 생성형AI를 도입한 기업들은 생성형AI 개발과 배치를 가로막는 가장 큰 장애물로 AI 인력 부족과 실행 장애를 제치고 규제 컴플라이언스를 꼽았다.[39] 규제 불확실성, 새로운 규제의 범위와 복잡성은 분명 큰 장애물이다. 에이전틱 AI 솔루션은 기업들의 규제에 대한 이해와 컴플라이언스를 지원해, 생성형AI 도입을 한층 가속화하는 데 도움이 될 수 있다.

- **에이전트 구축 및 조율:** 에이전틱 AI는 범산업적 또는 산업 특화 자동화 솔루션으로 부상하고 있다. 하지만 기업들은 관련 시장이 형성될 때까지 기다릴 필요 없이 자체 에이전트와 멀티에이전트 시스템을 구축할 수 있다. 구글Google의 생성형AI 모델 프로토타입을 제작 및 테스트할 수 있는 노코드no-code 툴인 버텍스Vertex를 활용하면, 이전 마케팅 캠페인에 기반해 마케팅 자료를 생산하는 등 특정 태스크를 위한 에이전트를 만들 수 있다.[40] 언어모델 기반 오픈소스 프레임워크인 랭체인LangChain을 활용하면 멀티에이전틱 시스템을 구축할 수 있다. 예를 들어, 스타트업 패러다임Paradigm은 다수의 에이전틱 AI가 협업해 다양한 출처에서 데이터를 수집해 정형화하고 태스크를 완수하는 '스마트 스프레드시트'smart spreadsheet를 출시했다.[41]

결론

에이전틱 AI는 업무흐름 전반 또는 일부를 자동화함으로써 지식 근로자의 생산성을 개선할 막대한 잠재력을 지니고 있다. 싱글 에이전트로서 또는 여타 에이전트와 협업해 독립적으로 행동할 수 있는 능력은 오늘날의 챗봇이나 코파일럿과 차별화된다. 하지만 에이전틱 AI는 아직 개발과 도입의 초기 단계이다. 초기 활용 사례들이 인상적이기는 하지만, 에이전틱 AI 또한 실수를 하거나 오류의 반복 회로에 갇힐 수 있다. 멀티에이전트 시스템에서는 '환각' 현상이 에이전트 사이에서 전염될 수 있다. 다른 에이전트를 잘못된 단계로 유도해 사실과 다른 답을 내놓는 것이다.[42] 에이전틱 AI는 주로 자율적으로 태스크를 수행하지만, 휴먼인더루프HITL, human in the loop*보다 덜 제한적인 휴먼온더루프HOTL, human on the loop** 방식으로 에이전틱 AI의 결정을 사람이 리뷰하는 것이 오늘날 더 현실적인 에이전틱 AI 배치 방식이 될 수 있다. 생성형AI 에이전트가 태스크 수행 과정에서 벽에 부딪쳤을 때 사람 전문가에게 도움을 구해 문제를 해결하고 남은 단계를 수행할 수 있다. 이러한 모델에서 에이전틱 AI는 가치 있는 업무를 수행하면서 경험으로부터 학습할 수 있는 신입 직원과도 같다.[43]

* 휴먼인더루프(HITL, human in the loop)는 AI가 도출한 데이터를 실제로 중요한 비즈니스 프로세스에 적용하기 전에 인간이 데이터를 직접 검증 및 수정해 생성물의 품질을 보장하는 방식을 뜻한다.
** 휴먼온더루프(HOTL, human on the loop)는 자동화 시스템이 대부분의 태스크를 독립적으로 수행하되, 인간이 시스템을 모니터링하고 필요한 경우 개입하는 방식을 뜻한다. HITL과 달리 자동화를 기본 시스템으로 하되, 특정 상황에서만 인간이 개입한다.

일부 기업들이 일관적이고 신뢰할 수 있는 에이전틱 AI를 개발하기 위해 수십 억 달러를 투자하고 있지만, 언제 어떠한 환경에서 성공할지 알 수 없다. 에이전틱 AI의 광범위한 도입이 당장 2025년에 가능할 수도 있고 5년 후에나 가능할 수도 있다. 더불어 또 다른 돌파구가 될 혁신이 필요할 수도 있고 기존의 AI 기술과 훈련 방식을 조금 수정하는 것만으로도 가능할 수 있다. 현재 에이전틱 AI를 개발하는 빅테크 기업들과 스타트업들이 성공한다면 게임의 법칙이 단숨에 바뀔 것이다. 자동화된 생성형AI 에이전트가 멀티모달 데이터를 처리하고, 툴을 사용하고 여타 에이전트를 조율하며, 과거의 경험을 기억하고 학습하며, 일관적이고 신뢰할 수 있는 방식으로 태스크를 수행할 수 있다. 더 나아가 기업들이 노코드 환경에서 단기 대화형 텍스트 프롬프트만으로도 맞춤형 생성형AI 에이전트를 손 쉽고 빠르게 만들 수도 있다.

에이전틱 AI의 미래는 이처럼 강력하고 관련 기술이 빠르게 발전하는 만큼, 기업들은 다음의 접근법을 고려해 지금 당장 준비를 시작할 필요가 있다.

- **에이전틱 AI를 위한 업무 흐름의 우선순위 정립과 재설계:** 에이전틱 AI의 역량과 조직에 가장 높은 가치를 창출할 수 있는 부분에 기반해 에이전틱 AI가 수행하기에 가장 적합한 태스크와 업무 흐름을 파악하라. 불필요한 단계를 없애기 위한 업무 흐름의 재설계도 필요하다. 또한 에이전틱 AI 솔루션에게 명확한 목표를 제시하고,

태스크 수행에 필요한 데이터와 툴, 시스템에 접근할 수 있도록 하는 것도 중요하다. 에이전틱 AI는 여타 에이전트가 환경에 적응하도록 도울 수 있지만, 프로세스가 정리되지 않은 채 최적화돼 있지 않으면 실망스러운 결과를 도출할 수 있다.

- **데이터 거버넌스와 사이버 보안에 주력:** 에이전틱 AI로 가치를 창출하려면, 에이전틱 AI가 민감하지만 가치 있는 기업 데이터와 더불어 내부 시스템과 외부 자원 등에 모두 접근할 수 있어야 한다. 따라서 에이전틱 AI를 도입하기에 앞서 기업들은 엄격한 데이터 거버넌스와 사이버 보안을 마련해야 한다. 최근 서베이에 따르면, 생성형AI를 선도적으로 도입하는 기업들이 가장 주력하는 IT 투자는 데이터 관리(75%)와 사이버 보안(73%)이다.[44] 이처럼 집중적 투자에도 불구하고, 응답자의 58%는 AI 모델에 민감한 데이터를 활용하는 것과 데이터 보안 관리에 깊은 우려를 나타냈다. 또한 생성형AI 리스크와 거버넌스 관리에 충분히 준비가 돼 있다는 응답자는 23%에 그쳤다. 다시 말해 선도적으로 생성형AI를 도입하는 기업들조차 에이전틱 AI에 대한 대비가 돼 있지 않다는 의미다.

- **리스크와 보상 간 균형:** 에이전틱 AI를 처음 도입하는 기업들은 에이전틱 AI에게 허용할 수 있는 자율성과 데이터 접근성의 수준을 결정해야 한다. 처음에는 사람의 감독 하에 중요도가 덜한 데이터에 기반한 리스크가 낮은 활용 사례에 에이전틱 AI를 적용해, 안전한 에이전틱 AI 활용을 위한 데이터 관리와 사이버 보안, 거버넌스를 구축해 나가는 것이 바람직하다. 이러한 안전망이 자리를 잡은 후에는 전략적 데이터에 기반하여 더 많은 툴과 자율성을 발휘할 수 있는 태스크에 에이전틱 AI를 적용할 수 있다.

- **건전한 회의적 시각 유지:** 에이전틱 AI는 2025년 한 해 괄목할 만한 발전을 이뤄 범산업적 솔루션 및 산업 특화 솔루션으로 활용되기 시작할 수 있다. 인상적인 시연과 시뮬레이션, 제품 공개가 2025년 내내 이어질 수 있다. 하지만 앞서 지적한 문제와 장애물은 발전 속도만큼 빠르게 해결하기 어렵다. 이러한 문제들을 해결할 때까지, 에이전틱 AI는 통제된 환경에서만 적용돼 실질적인 기업 성과 개선을 이루기 어려울 수 있다. 건전하지만 신중한 접근 방식이 필요하다.

05
딥페이크와 사이버 보안, 쫓고 쫓기는 싸움

> 가짜 콘텐츠를 탐지하고 대응하는 노력이 강화되면서, 소비자와 창작자, 광고주들이 신뢰할 수 있는 인터넷 환경을 유지하기 위한 비용 부담을 안게 될 수 있다.

진짜처럼 보이지만 인공지능AI 툴로 만든 사진과 비디오, 오디오를 뜻하는 딥페이크deepfake가 확산되면서 온라인 콘텐츠에 대한 신뢰가 추락하고 있다. AI가 만든 콘텐츠가 많아지고 한층 정교화되면서, 온라인 상의 이미지와 비디오, 오디오가 잘못된 정보를 퍼뜨리고 사기행위를 저지르는 데 악용될 수 있다. 소셜미디어에는 그러한 합성 콘텐츠가 넘쳐나, 온라인 콘텐츠를 믿지 못하고 데이터가 악용될 것을 우려하는 사용자가 증가하고 있다.[1]

딜로이트의 2024년 서베이에 따르면, 소비자의 절반은 온라인 상 정보의 정확성과 신뢰성에 대해 1년 전보다 회의적 시각이 심화됐다고 답했다. 생성형AI에 익숙하거나 사용 중이라는 응답자 중 68%는 합성 콘텐츠가 속임수나 사기에 악용될 수 있다고 우려했고, 59%는 사람이 만든 미디어와 AI의 창작물을 구별하기 어렵다고 답했다. 생성형AI에 익숙하다는 응답자 중 84%는 생성형AI가 만든 콘텐츠에는 항상 분명한 라벨링labeling이 돼야 한다는 데 동의했다.[2]

라벨링은 미디어와 소셜미디어가 사용자들에게 합성 콘텐츠를 표시하는 한 가지 방안이 될 수 있다. 하지만 합성 콘텐츠를 만들고 기존 미디어를 조작할 수 있는 첨단 AI 모델이 딥페이크 기술에 적용되면서, 가짜를 탐지해내고 신뢰를 회복하기 위해 보다 복잡한 대응조치가 필요한 실정이다.

애널리스트들은 테크, 미디어, 소셜네트워크 부문 대기업들이 실행하는 딥페이크 탐지 관련 글로벌 시장 규모가 2023년 55억 달러에서 2026년 157억 달러로 연간 42% 증가할 것으로 내다봤다.[3] 딜로이트는 딥페이크 탐지 시장이 사이버 보안 시장과 비슷한 경로를 밟을 것으로 예상한다. 미디어와 테크 기업들은 콘텐츠 인증 솔루션과 컨소시엄에 투자함으로써 갈수록 정교해지는 가짜 콘텐츠에 한 발 앞서기 위한 노력을 지속할 것이다. 따라서 소비자와 창작자, 광고주들이 콘텐츠 신뢰성을 지키기 위해 증가하는 비용 부담을 안게 될 수 있다.[4]

가짜 콘텐츠 대응 노력은 크게 가짜 콘텐츠 탐지와 출처 시스템 구축의 두 가지 방향으로 이뤄지고 있다.

가짜 콘텐츠 탐지

테크 기업들은 딥러닝이나 컴퓨터 비전 등 방법으로 합성 미디어를 분석해 가짜나 조작된 콘텐츠를 찾아낸다. 머신러닝 모델을 이용해 딥페이크의 패턴과 비정상성을 찾아내는 것이다.[5] 또한 이러한 툴을 이용해 비디오나 오디오 콘텐츠에서 진짜 사람과는 다른 미묘한 입술의 움직임이라던가 목소리 톤의 변화와 같은 비일관성을 찾아낸다.[6]

일부 생성형AI 툴은 특정 콘텐츠가 생성형AI의 도움을 받아 만들어진 것인지 알아낼 수 있지만, 다른 모델이 만든 딥페이크는 탐지하기 어렵다.[7] 일부 가짜 탐지 툴은 조작된 증거나 생성형AI의 '흔적'을 찾기도 한다.[8] 신뢰할 수 있는 출처나 이미 알려진 가짜 콘텐츠의 특징에 기반해 '화이트리스트'whitelist와 '블랙리스트'blacklist를 구분하는 방식도 있다. 자연스러운 혈류, 얼굴 표정, 어조 등 인공과 대비되는 인간만의 특징을 찾는 툴도 있다.[9]

현재 딥페이크 탐지 툴의 정확도는 90%를 넘는다.[10] 하지만 오픈소스 생성형AI 모델을 이용해 합성 미디어를 만들면 이러한 툴로도 딥페이크를 탐지하지 못할 수 있다. 콘텐츠 생성을 자동화하는 기술이 현재의 탐지 기술을 능가할 수 있으며, 생성형AI 툴이 사용자 프롬프트에 따라 생성물에 미묘한 수정을 가하면 가짜 콘텐츠를 교묘히 숨길 수 있다.[11]

소셜미디어 플랫폼도 AI 툴을 이용해 문제의 소지가 있는 이미지나 비디오 콘텐츠를 탐지한다. AI가 의심스러운 콘텐츠에 상대 평가를 내리면, 사람이 가장 수상한 콘텐츠를 검토해 최종적으로 딥페이크를 판별한다. 하지만 이러한 방식은 시간과 비용이 많이 소요되기 때문에, 머신러닝의 도움으로 프로세스의 시간을 줄이려는 노력이 이뤄지고 있다.[12]

이는 사이버 보안 시장이 형성된 경로와 흡사한 양상이다. 보안에 주력하는 기업들이 데이터와 네트워크 보호에 계층적 접근법을 도입했던 것처럼, 미디어와 소셜미디어 기업들도 디지털 콘텐츠의 진위를 파악하기 위해 콘텐츠 출처 시스템과 더불어 다양한 툴을 활용할 가능성이 크다.

출처와 신뢰 시스템 구축

일부 기업들은 출처와 모든 수정 내역을 표시한 암호화 메타데이터cryptographic metadata* 즉 디지털 워터마크digital watermark**를 미디어 파일에 추가하는 방법을 모색하고 있다.[13]

* 암호화 메타데이터(cryptographic metadata)는 디지털 콘텐츠와 데이터의 출처 확인, 무결성 검증, 통신의 보안, 접근 권한 관리 등을 위해 사용되는 암호화 관련 정보를 뜻한다.
** 디지털 워터마크(digital watermark)는 이미지, 비디오, 오디오, 문서 등 디지털 콘텐츠의 소유권, 무결성, 출처 확인 등을 보장하기 위해 눈에 띄지 않는 형태로 삽입된 데이터를 뜻한다. 콘텐츠의 무단 복제나 조작을 방지하고, 저작권 보호 및 보안 목적으로 사용된다.

소셜플랫폼은 미디어 매체, 기기 제조사, 테크 기업들과 산업간 컨소시엄을 맺어 콘텐츠 진위 검증을 위한 영구적 표준 마련 노력을 펼치고 있다.

일례로, 딜로이트와 더불어 다양한 테크 및 미디어 기업들은 콘텐츠 출처 및 진위 확인을 위한 연합C2PA, Coalition for Content Provenance and Authenticity에 가입해, AI가 생성한 이미지를 더욱 손쉽게 확인할 수 있는 C2PA 메타데이터 표준을 사용하기로 했다.[14] C2PA의 메타데이터 기술은 최초 창작부터 모든 편집 과정까지 이미지 생성의 전全 주기를 상세 로그로 기록한다.[15] C2PA 기록을 해독하면 콘텐츠 매체와 사용자들은 해당 이미지의 출처와 진위 여부를 확인할 수 있다.

일부 소셜플랫폼은 창작자들을 대상으로 검증 옵션 시스템을 마련하기도 했다. 가짜가 아님을 증명하는 검증서를 제출하는 창작자들에게 수수료를 지불하는 방식이다. 창작자들을 수익 공유 프로그램에 참여토록 해 검증 활동에 인센티브를 제공할 수도 있다.[16]

각종 온라인 채널에서 AI 콘텐츠가 확산되는 만큼, 계정의 운영자가 사람인지 AI인지를 파악하는 것도 신뢰 악화를 막는 데 도움이 된다.[17] 다만 이에 따른 비용 부담을 창작자와 광고주, 사용자들에게 계속 전가할 수 있는지는 고민해야 할 부분이다.

딥페이크 대응 규제

일부 국가에서 콘텐츠 진위 관련 규제 조치가 마련됐지만,[18] 더욱 포괄적인 글로벌 규제가 필요하다. 딥페이크의 위험성에 대한 대중의 인식을 제고하는 것도 중요하다. 이를 위해 대중이 가짜와 진짜 미디어를 구분할 수 있는 정보를 제공할 필요가 있다.

미국에서는 연방 상원 상업·과학·운송 위원회가 AI 생성 콘텐츠에 디지털 워터마크를 의무화하는 법안을 검토 중이다.[19] 또 캘리포니아 주 정부는 AI 콘텐츠의 라벨링을 의무화하는 법안 'AB 3211'을 상정했다. 이에 따르면, 기기 제조사들은 사진에 출처 메타데이터가 표시되도록 기기를 업데이트해야 하고 온라인 플랫폼들은 온라인 콘텐츠의 출처 메타데이터를 공개해야 한다. 법안이 통과되면 2026년부터 실효될 예정이다.[20] 여타 주 정부도 잘못된 정보를 퍼뜨리려는 의도로 동의 없이 만들어진 딥페이크의 제작과 배포를 불법으로 규정하는 법률을 제정했다.[21] 미국 연방거래위원회FTC는 개인을 모방한 딥페이크의 창작과 유포를 금지하는 새로운 규제를 마련 중이다.[22]

유럽연합EU은 인공지능법AI Act 수정안에서 투명성을 강조하며 AI 생성 콘텐츠와 딥페이크의 라벨링 의무를 명시했다. 이는 AI 기술의 발전을 지원하면서도 사용자들이 콘텐츠의 진위를 파악할 수 있도록 하기 위함이다. 또 EU 집행위원회는 AI의 발전과 활용을 지원하고 AI가 창작 또는 조작한 콘텐츠의 효과적인 라벨링을 장려하기 위해 AI 사무국AI Office을 수립했다.[23]

딥페이크 기술의 급격한 발전 속도에 발 맞춰 유연하고 발 빠르게 대응할 수 있는 규제 프레임워크가 시급하다.

결론

미디어 매체와 소셜플랫폼은 사진과 비디오, 오디오 클립의 진위를 파악하기 위해 콘텐츠 분석과 출처 확인 두 가지 기술을 개발하는 데 투자를 지속할 것이다. 생성형AI를 활용해 더욱 많은 합성 미디어가 확산되고 딥페이크 생성 모델과 탐지 회피 기술도 빠르게 발전하는 만큼, 이러한 투자는 갈수록 중요해질 것이다.

특히 생성형AI의 능력이 한층 강력해지고 다각화되고 있으므로, 딥페이크 기술보다 한 발 앞서는 것이 중요하다. 최근에는 혈액량 탐지 및 안면 분석 등의 첨단 기술이 진짜와 가짜를 구분하는 데 활용되고 있다. 하지만 사이버 보안 툴과 마찬가지로 이러한 딥페이크 탐지 기술이 최종 사용자와 소비자들에게 불편을 초래해서는 안 된다. 사용자 경험을 침해하지 않는 최선의 방식으로 콘텐츠의 무결성을 보장할 필요가 있다. 현재로서는 사용자 경험에 부정적 영향을 주지도 않고 분석을 위한 실시간 컴퓨팅 주기가 없어도 되는 디지털 워터마크가 가장 효율적 수단이다.[24]

딥페이크 탐지에 머신러닝 모델이나 외부 업체를 이용하는 기업들은 이미지와 비디오, 오디오를 아우르는 다양한 고품질 데이터 세트를 활용하는 툴과 업체를 우선적으로 선택할 필요가 있다. 또한 이러한 데이터 세트는 다양한 인구층을 통합해 공정성을 보장하고 편향성을 최소화하는 것이 중요하다.[25]

테크 및 미디어 기업들은 범산업적 협력을 통해 딥페이크 탐지와 콘텐츠 진위 검증을 위한 표준 수립 노력을 기울여야 한다.[26] 일례로, 기기 제조사와 미디어 기업이 창작물과 발행물의 진위를 공동 검증하는 경우 디지털 워터마크가 여타 방식보다 효과적일 수 있다. 이처럼 공동의 노력이 펼쳐지면 더욱 강력하고 포괄적인 기준이 수립돼 디지털 콘텐츠 전반의 보안과 신뢰가 개선될 것이다.

각 산업의 기업들은 생성형AI로 더욱 위협적인 소셜 엔지니어링 공격이 가능해져 일부 가짜 탐지 조치를 무력화할 수 있음을 염두해야 한다.[27] 따라서 비디오와 오디오 기반 프로세스의 경우 추가 검증 계층을 이행할 필요가 있다. 최종 사용자들 사이에서도 신뢰할 수 있는 출처에 기반해 정보를 대조 검토하고 다중 인증 방식을 이용해 딥페이크 관련 리스크를 완화하는 노력이 자리잡아야 한다. 따라서 기업들은 계속 진화하는 딥페이크 기술에 발 맞춰 사용자 대상 교육을 지속하는 것이 중요하다.

테크 및 미디어 기업들이 이러한 전략을 충실히 이행하면 딥페이크 기술이 가하는 위협을 방지함과 동시에 디지털 콘텐츠의 무결성과 신뢰를 유지하는 리더로서의 역할을 다질 수 있다. 2025년은 디지털 세계의 불확실성이 증대하는 가운데 기업들이 신뢰할 수 있는 콘텐츠 환경을 만들어 스스로 신뢰할 수 있는 콘텐츠 출처로 자리매김하는 중요한 기로가 될 수 있다.