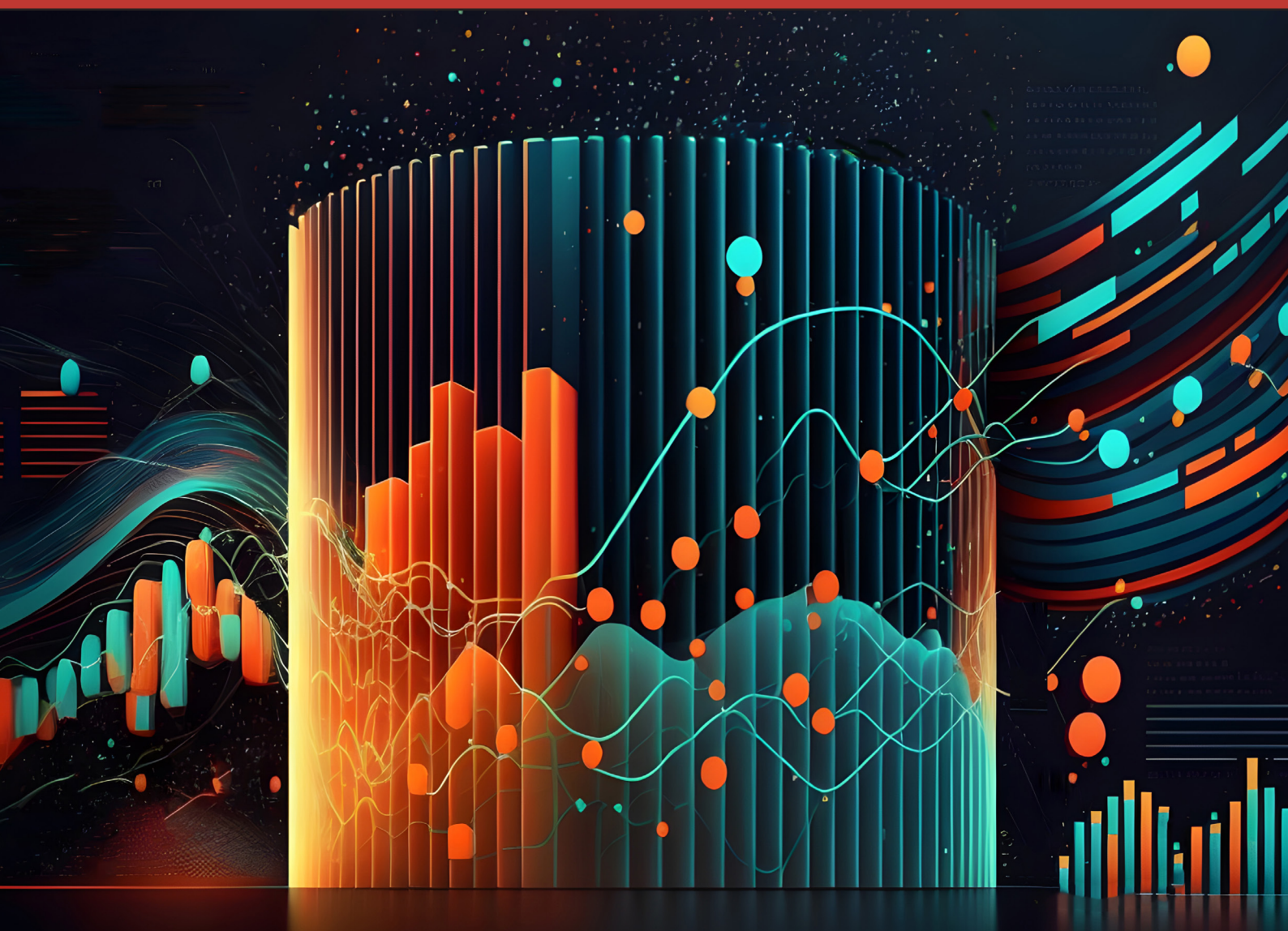


데이터와 AI가 주도하는 성장을 위한 기반 마련



서문

“데이터와 AI가 주도하는 성장을 위한 기반 마련”은 Databricks의 후원으로 작성된 MIT Technology Review Insights 보고서입니다. 본 보고서를 작성하기 위해 MIT Technology Review Insights는 고위급 데이터 및 기술 경영진을 대상으로 글로벌 설문조사를 실시했습니다. 또한 본 보고서는 2023년 6월~8월에 대규모 민간 및 공공 부문 조직의 CIO, CTO 및 CDO와 진행한 심층 인터뷰를 바탕으로 작성되었습니다.

본 보고서의 저자는 Denis McCauley, 편집자는 Teresa Elsey, 발행인은 Nicola Crepaldi입니다. 본 연구는 편집의 독자성을 보유하며, MIT Technology Review Insights의 관점이 반영되었습니다.

소중한 시간을 내어 의견을 주신 다음 경영진들께도 감사드립니다.

Jack Berkowitz, 최고 데이터 책임자, ADP

Sanjay Bhakta, 최고 제품 및 기술 책임자, Condé Nast

Murali Brahmedesam, 최고 기술 책임자 겸 엔지니어링 책임자, Razorpay

Jon Francis, 최고 데이터 및 분석 책임자, General Motors

Deb Hall Lefevre, 영업 담당 부사장 겸 최고 기술 책임자, Starbucks

Yemi Oshinnaiye, 최고 정보 책임자, U.S. Transportation Security Administration

Jeffrey Reid, 최고 데이터 책임자, Regeneron Genetics Center

John Roese, 글로벌 최고 기술 책임자, Dell Technologies

Naveen Zutshi, 최고 정보 책임자, Databricks

설문조사 정보

본 보고서의 기반이 되는 설문조사는 2023년 6월~8월에 MIT Technology Review Insights에 의해 진행되었으며, 전 세계 600명의 고위 기술 경영진에게서 설문 응답을 받았습니다.

설문조사 응답자 전원은 북미, 유럽, 아시아 태평양, 중동 지역 12개국에 본사를 둔 대기업 또는 공공 부문 조직에 속해 있습니다. 약 75%는 최고 경영진이며, 나머지는 SVP, VP, IT/AI/데이터/엔지니어링 책임자 및 유사한 역할로 구성되어 있습니다. 설문조사에 참여한 기업들은 8개의 산업 부문에 속하며, 모두 연간 수익이 5억 달러가 넘어가는 조직입니다.

직책

- 최고 정보 책임자
- 최고 기술 책임자
- 최고 데이터/분석 책임자
- 최고 데이터 사이언티스트
- 최고 엔터프라이즈/데이터 아키텍처 설계자
- SVP, VP, IT/AI/데이터/엔지니어링 책임자, 또는 유사한 역할

8개 산업 부문

- 리테일 및 CPG
- 미디어 및 엔터테인먼트
- 통신
- 의료 및 생명 과학
- 금융 서비스
- 에너지
- 제조
- 정부 및 공공 부문

응답자는 12개 국가 출신이며 4개 지역을 대표

북미

- 캐나다
- 미국

유럽

- 프랑스
- 독일
- 네덜란드
- 영국

아시아 태평양

- 오스트레일리아
- 인도
- 일본
- 싱가포르
- 대한민국

중동

- 이스라엘

목차

01 핵심 요약 4

02 AI 주도 성장을 위한 새로운 전망 6

03 현대화 및 간소화 10

04 생성형 AI 가속기 14

05 시티즌(citizen) 개발자의 부상 20

06 결론.....23



01

핵심 요약

기업의 AI 도입은 한 단계 더 도약할 준비가 되어 있습니다. 생성형 AI의 기능은 전사적으로 경영진의 관심을 끌었으며 기술 경영진은 이를 배포하거나 실험하기 위해 빠르게 움직이고 있습니다. 많은 조직에서는 내년에 더 폭넓은 AI 기능 제품군과 이를 지원하는 데이터 인프라에 대한 지출을 두 자릿수로 늘릴 계획입니다. 불리한 경제 상황에 대한 우려에도 불구하고 경영진은 데이터와 AI를 활용하여 매출과 수익 측면에서 조직이 더 큰 성장을 이룰 수 있는 기회를 모색합니다.

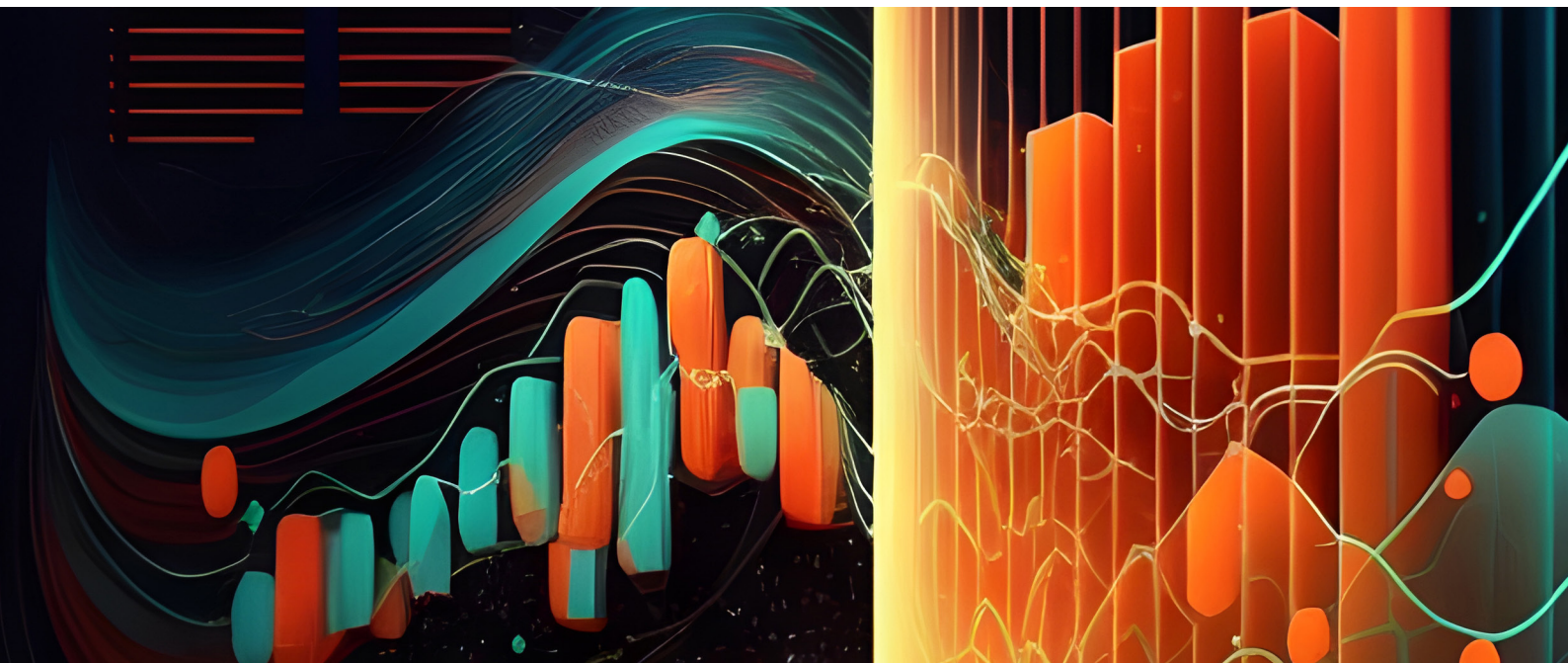
600명의 기술 리더를 대상으로 실시한 글로벌 설문조사와 심층 인터뷰를 바탕으로, 본 보고서에서는 조직이 데이터가 AI가 주도하는 미래를 재편성하는 데 중점을 두고 있는 것으로 나타났습니다. 기술 경영진이 효율성을 높이고 새로운 성장 영역을 찾기 위해 노력함에 따라 데이터 아키텍처부터 AI 지원 자동화에 이르기까지 모든 주제가

논의되고 있습니다. 이와 동시에 데이터와 AI의 강력한 기능을 보편화해야 한다는 압박감으로 인해 데이터 거버넌스와 보안을 강화해야 한다는 급박한 분위기가 조성되었습니다.

주요 연구 결과는 다음과 같습니다.

• **CIO는 데이터와 AI에 대한 투자를 늘리고 있습니다.**

잠재고객의 기대치는 점점 커지면서, 새로운 경쟁 압력, 어려운 경제 여건, 전례 없는 혁신 속도에 직면한 기술 리더들은 이전보다 더 큰 성장을 이루기 위해 데이터와 AI 자산이 필요합니다. 그들은 이러한 미래를 실현하기 위해 투자하고 있습니다. 설문조사에 참여한 모든 조직은 내년에 데이터 인프라를 현대화하고 AI를 도입하는 것에 대한 지출을 늘리고, 거의 절반(46%)은 증가율이 25%를 초과할 것이라고 밝혔습니다.



• **데이터와 AI 시스템의 통합이 무엇보다 중요합니다.**

데이터 및 AI 시스템의 확산은 특히 설문조사에 참여한 연간 수익이 100억 달러 이상인 조직에서 광범위하게 나타납니다. 그 중 81%는 이러한 시스템을 10개 이상 운영하고, 28%는 20개 이상을 사용하고 있습니다. 인터뷰에 참가한 경영진의 목표는 여러 시스템을 조금씩 줄여 전사적으로 데이터를 통합 플랫폼에 연결하여 사일로를 해소하고 AI 이니셔티브의 확장을 지원하는 것입니다.

• **AI의 보편화는 거버넌스의 중요성을 한층 높였습니다.**

여러 비즈니스 부서와 직원들이 생성형 AI 사용을 요구함에 따라 경영진들은 이 기술에 대한 거버넌스 프레임워크가 필요한 데이터 정확성과 무결성뿐만 아니라 적절한 데이터 개인정보 보호 및 보안도 제공할 수 있다는 확신을 추구합니다. 이것이 바로 응답자의 60%가 데이터와 AI에 대한 단일 거버넌스 모델이 “매우 중요하다”고 말하는 이유일 것입니다.

• **경영진은 AI 도입이 단기적으로 혁신을 가져올 것으로 기대합니다.**

설문조사 응답자의 81%는 AI가 향후 2년 동안 업계의 효율성을 최소 25% 높일 것으로 기대하고, 1/3은 이익이 최소 50%가 될 것이라고 말합니다.

• **생성형 AI가 확산되면서 유연한 접근 방식이 선호되고 있습니다.** 88%의 조직이 생성형 AI를 사용 중이며, 이 중 약 1/4(26%)은 이미 생성형 AI에 투자하여 도입했고 나머지 62%는 실험 단계에 있습니다. 58%은 이러한 기능을 개발하기 위해 하이브리드 접근 방식을 취하고 있는데, 일부 사용 사례에 벤더의 대규모 언어 모델(LLM)을 사용하고 IP 소유권, 개인정보보호, 보안 및 정확성이 엄격하게 요구되는 경우에는 자체 모델을 구축하고 있습니다.

• **레이크하우스는 생성형 AI 시대를 위한 데이터 아키텍처가 되었습니다.**

설문조사에 참여한 조직 중 거의 3/4이 레이크하우스 아키텍처를 도입했고 나머지 대부분은 향후 3년 내에 그렇게 할 것으로 예상하고 있습니다. 설문조사 응답자들은 실시간 분석(72%), 최신 기술의 간편한 통합(66%) 및 플랫폼 간 라이브 데이터 공유(64%)를 위한 스트리밍 데이터 워크로드를 지원하기 위해 데이터 아키텍처가 필요하다고 말합니다. 레이크하우스 도입업체의 99%는 아키텍처가 데이터 및 AI 목표 달성에 도움이 된다고 말했으며, 74%는 그 도움이 “중요하다”도 말했습니다.

• **인력에 투자하면 데이터와 AI에서 더 많은 가치를 창출할 수 있습니다.**

본 설문조사에서 인재와 기술 격차가 조직의 다른 데이터 및 AI 과제를 무색케 합니다. 회사의 데이터 전략에서 개선이 필요한 부분을 묻는 질문에 응답자 중 가장 많은 비율(39%)이 인재에 대한 투자라고 답했습니다. 데이터 및 AI 플랫폼과 관련하여 직면하는 가장 큰 어려움은 이를 사용할 수 있도록 직원을 교육하고 역량을 강화하는 것으로, 40%가 이를 가장 큰 우려사항으로 꼽았습니다.

후속 보고서에서는 금융 서비스, 의료 및 생명 과학, 소매 및 소비자용 포장 식품, 제조, 미디어 및 엔터테인먼트, 정부 등 6개 부문에 걸친 추가 경영진 인터뷰를 통해 얻은 인사이트와 함께 이러한 설문조사 결과를 자세히 검토할 것입니다.

잠재고객의 기대치는 점점 커지면서, 새로운 경쟁 압력, 어려운 경제 여건, 전례 없는 혁신 속도에 직면한 기술 리더들은 이전보다 더 큰 성장을 이루기 위해 데이터와 AI 자산이 필요합니다.

02

AI 주도 성장을
위한 새로운 전망

거시 경제 환경의 불확실성 속에서 고위 경영자들은 데이터 및 AI를 포함한 기술 투자가 그 어느 때보다 조직의 효율성을 개선하고 조직에 더 큰 가치를 제공할 것으로 요구하고 있습니다.

생성형 AI의 출현은 기술 경영진에게 이를 실현하기 위한 도구를 추가로 제공합니다. 동시에 생성형 AI는 이러한 문제의 복잡성을 가중시킵니다. 기업의 많은 부분에서 이를 실험하기를 원하면서 CIO는 이러한 발전으로 인해 발생하는 막대한 데이터 처리 요구와 거버넌스 문제에 대처할 수 있을 만큼 충분히 데이터 인프라가 견고한지 확인해야 합니다.

본 연구에 참여한 기술 리더들은 이러한 문제를 AI가 매출과 수익 측면에서 조직이 상당한 성장을 실현할 수 있는 기회로 보고 있습니다. 설문조사 응답자의 70%가 AI 프로젝트가 비용 절감에 도움이 되는 것이 매우 중요하다고 답했고, 같은 비율의 응답자가 이러한 프로젝트가 새로운 수익의 원천이 된다는 것이 매우 중요하다고 답했습니다(단, 산업과 지역에 따라 차이가 있음). 경영자들이 현 시대를 위기에 대비할 시기로 생각하지 않는 것은 분명합니다.

스타벅스의 부사장이자 최고 기술 책임자인 Deb Hall Lefevre에 따르면 AI는 “커피 원산지부터 매장까지(From Bean To Cup)” 비즈니스 전반에서 경험을 강화할 것입니다. 예를 들어, Hall Lefevre는 회사의 연구원은 물론 매장의 바리스타가 생성형 AI를 사용하여 새로운 맞춤형 음료를 발견 및 개발하고 드라이브 스루

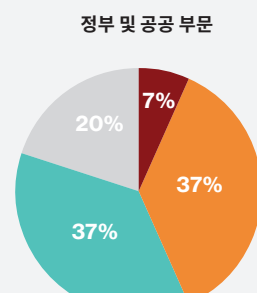
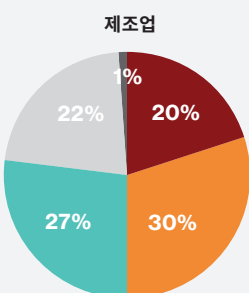
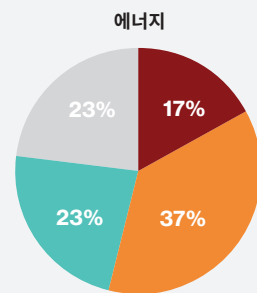
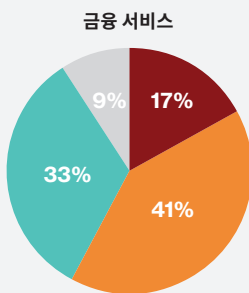
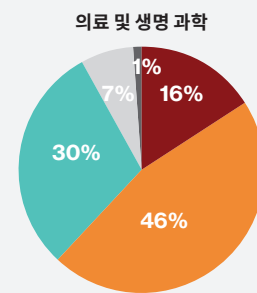
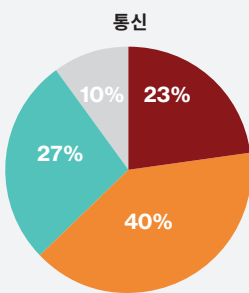
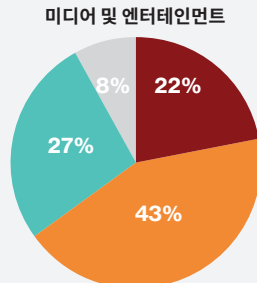
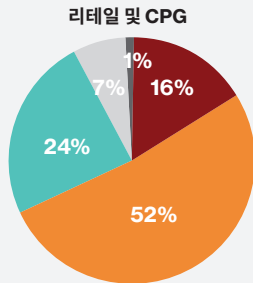
“발전하고 있는 AI와 머신 러닝에 투자한 덕분에 몇 년 전보다 더 효과적으로 효율성을 높일 수 있게 되었습니다. 이를 위해 백오피스에는 HR 챗봇을 활용하고, 공장 현장에는 예측적 유지관리를 적용하며, IT 운영 측면에서는 소프트웨어 개발을 확장하고 프로덕션화할 것입니다.”

Jon Francis, 최고 데이터 및 분석 책임자, General Motors

그림 1: 도입 속도

귀사가 속한 산업이나 분야에서 현재 AI를 도입하는 속도를 어떻게 평가하시나요?

■ 매우 빠름 ■ 빠름 ■ 보통 ■ 느림 ■ 매우 느림



출처: MIT Technology Review Insights 설문조사, 2023

경험에서 빚어지는 마찰을 줄이는 방법을 찾을 것으로 기대합니다. Hall Lefevre는 이렇게 말합니다. “스타벅스의 차별점은 규모입니다. 대규모 혁신을 신속하게 추진할 수 있으면 실제로 경쟁 우위를 구축할 수 있습니다.”

대체로 기술 리더들은 더 적은 비용으로 더 많은 성과를 내도록 요구받지 않습니다. Gartner는 전 세계 IT 지출이 2023년에 4.3%, 2024년에 8.8% 증가할 것으로 전망합니다.¹ 이러한 성장의 대부분은 데이터 및 AI에 대한 지출을 포함하는 소프트웨어 범주에 집중될 것입니다. 실제로 설문조사에 참여한 모든 경영진은 내년엔 데이터 인프라의 현대화와 AI 도입에 대한 조직의 지출이 증가할 것으로 예상했으며, 거의 절반(46%)은 지출이 25% 이상 증가할 것이라고 답했습니다.

낙관적인 전망일 수도 있지만 설문조사 응답자 중 대부분은 동종 업체의 유사한 조치를 인지하고 있습니다. 설문조사 응답자 중 60%는 업계의 AI 도입이 빠르거나 “매우 빠르다”고 말합니다(그림 1 참조). 역학은 산업별로 다소 다르지만(자세한 내용은 후속 산업 보고서 참조), 대체로 기업의 AI 도입이 빠른 속도로 진행되고 있으며 이를 지원하기 위한 데이터 현대화가 최우선 과제로 남아 있습니다.

AI 활용에 따른 효율성 향상으로 비즈니스 성장 보장

설문조사에 참여한 경영진은 AI가 향후 2년 동안 업계에 가져올 효율성 향상에 대해 확실히 낙관하고 있습니다. 81%는 해당 기간 동안 효율성이 25% 이상 향상될 것으로 기대하고, 33%는 50%를 초과할 것이라고 생각합니다(그림 2 참조).

인터뷰에 참여한 경영진은 AI가 주도하는 성장의 상당 부분이 효율성 향상에서 비롯될 수 있다는 점에 동의합니다. 자동차 제조업체 General Motors(GM)의 최고 데이터 및 분석 책임자인 Jon Francis는 “발전하고 있는 AI와 머신 러닝에 투자한 덕분에 몇 년 전보다 더 효과적으로 효율성을 높일 수 있게 되었습니다.”라고 말합니다. “이를 위해 백오피스에는 HR 챗봇을 활용하고, 공장 현장에는 예측적 유지관리를 적용하며, IT 운영 측면에서는 소프트웨어 개발을 확장하고 프로덕션화할 것입니다.”

일부 조직에서는 효율성을 한층 더 높이기 위해 AI 지원 자동화를 사용하는 것을 목표로 합니다. 미국 교통안전국(US Transportation Security Administration, TSA)의 최고 정보 책임자인 Yemi Oshinnaiye는 “자가 학습 알고리즘의 도움으로 핵심 프로세스를 자동화할 수 있으면 인력을 늘리지 않고도 더 빠르게 더 많은 성과를 실현할 수 있게 될 것입니다.”라고 덧붙입니다. “궁극적으로 대규모로 빠르게 자동화를 실현할 수 있습니다.”

가속화된 AI의 우선순위를 엄격하게 지정

조직은 AI를 도입하기를 바라지만, 오늘날 기술 리더는 AI 프로젝트 포트폴리오에 투자하고 이를 관리하는 방법을 매우 신중하게 결정할 가능성이 높습니다. Dell Technologies의 글로벌 최고 기술 책임자인 John Roesse는 이렇게 말합니다. “현 상황에서는 리소스를 보존하는 데 주의를 기울여야 합니다. 그렇다고 AI와 고급 데이터 기술의 도입을 멈출 수는 없습니다. 따라서 훨씬 더 구체적인 계획을 바탕으로 투자해야 합니다. 이러한 시기에는 자원이 한정되어 있기 때문에 어떤 프로젝트를 우선적으로 처리할지 선택해야 합니다.”

무엇보다도 AI 프로젝트의 우선순위를 지정하는 데 주력해야 하며, 이는 부분적으로 경제적 불확실성에 의해 좌우됩니다. Databricks의 최고 정보 책임자인 Naveen Zutshi는 이렇게 말합니다. “따라서 이전보다 훨씬 더 무자비하게 프로젝트의 우선순위를 정하게 되었습니다. 프로젝트 수를 줄였지만 회사에 상당한 영향을 미칠 수 있는 보다 장기적이고 획기적인 프로젝트를 추진하고 있습니다. IT 조직이 시간을 할애하는 경우가 많은 번거로운 작업을 모두 최소화했습니다.”

“그렇다고 AI와 고급 데이터 기술의 도입을 멈출 수는 없습니다. 따라서 훨씬 더 구체적인 계획을 바탕으로 투자해야 합니다.”

John Roesse, 글로벌 최고 기술 책임자,
Dell Technologies



Condé Nast: 소셜 미디어 최적화를 위한 AI

대형 웹 플랫폼은 콘텐츠 게시자를 힘들게 하고 있습니다. 이는 특히 뉴스 제공자에 해당되는 이야기인데, Facebook과 Google이 사이트에 포함하는 뉴스 링크의 양을 줄이면서 이에 대한 비용을 지불하기를 거부하고 있기 때문입니다. 그러나 Condé Nast와 같은 라이프스타일 전문 잡지의 출판사들도 이러한 플랫폼에서 콘텐츠를 게시할 공간을 확보하는 데 어려움을 겪고 있습니다. 이 회사의 최고 제품 및 기술 책임자인 Sanjay Bhakta는 이렇게 말합니다. “자체 플랫폼에서 고객과 상호작용하려는 기술 대기업이 점점 많아지고 있습니다. 그 결과 더 이상 예전처럼 사이트로 유입되는 트래픽의 양을 얻지 못합니다. 이제 목표는 기존 고객과 잠재 고객을 우리의 온라인 자산으로 유도하고 계속 유지하는 것입니다. 그리고 바로 여기서 콘텐츠로 수익을 창출합니다.”

Bhakta는 AI가 이러한 과제를 해결하는 데 필수 요소라고 말합니다. 예를 들어 그의 팀은 Condé Nast가 소셜 미디어에 게재하는 콘텐츠를 최적화하는 모델을 구축했습니다. Falcon이라고도 하는 이 모델은 사람들이 소셜 미디어에서 소비하는 콘텐츠를 분석한 다음 이를 회사의 자체 콘텐츠와

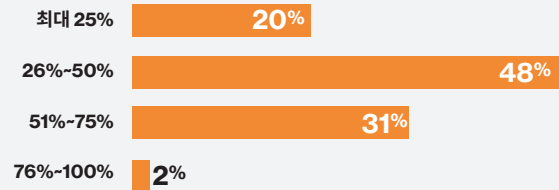
비교하여 소셜 미디어 사용자를 웹사이트로 유도할 가능성이 가장 높은 콘텐츠를 예측합니다. Bhakta는 모델에 필요한 대량의 데이터를 수집하고 훈련시키는 데 상당한 시간이 걸렸지만 모델이 회사의 사이트로 이동하는 방문자 수에 미치는 영향을 고려하면 그만한 가치가 있다고 말했습니다.

Bhakta는 이렇게 덧붙입니다. “Falcon은 소셜 미디어 팀이 게시해야 할 콘텐츠를 파악하는 데 놀랄만한 도움을 주었습니다. 우리는 그러한 결정을 내리는 인간과 비교하여 모델을 테스트했고, 언제나 머신이 추천하는 것이 항상 더 큰 영향을 미친다는 사실을 발견했습니다. 이 모델이 지속적으로 우리 웹사이트로 트래픽을 유도할 수 있다는 사실이 입증되었기 때문에 현재 적극적으로 사용하고 있습니다. 이 모델은 우리에게 매우 중요한 마케팅 도구가 되었습니다.”

생성형 AI의 도입으로 사업부와 부서 전반에 걸쳐 AI의 인기가 높아진 것도 프로젝트 우선순위 지정에 더 엄격한 규칙을 적용해야 하는 또 다른 이유입니다. Hall Lefevre에 따르면 새로운 지표와 프레임워크를 만드는 것보다 기존 지표와 프레임워크를 적용하는 데 더 엄격한 요구 사항이 필요합니다. 그녀는 이렇게 말합니다. “비즈니스의 모든 측면에서 사용 사례와 개념 증명이 쓰나미처럼 쏟아지고 있습니다. 그것도 과거의 어떤 새로운 기술 트렌드보다도 빠른 속도로요. 프레임워크 관점에서 우리는 예상되는 비즈니스 영향, 복잡성, 확장 가능성에 대한 테스트 비용에 대해 같은 질문을 하고 있습니다. 여기에서 핵심은 어떤 우선순위 프레임워크를 사용하든 올바른 것에 집중하도록 하는 것이 그 어느 때보다 중요하다는 것입니다.”

그림 2: AI에서 기대되는 효율성 향상

AI가 향후 2년 동안 귀사가 속한 산업에 가져올 효율성 개선 효과는 어느 정도일까요?



출처: MIT Technology Review Insights 설문조사, 2023

Dell Technologies: AI 사용 사례의 기준 강화

코로나19로 인한 이동 제한과 함께 기업의 디지털화가 가속화되면서 기술 기업들은 이전에 누렸던 것보다 새로운 AI 및 관련 기능을 구축하는 데 투자할 수 있는 더 많은 권한을 부여받았습니다. Dell Technologies의 글로벌 최고 기술 책임자인 John Roese는 그런 시대는 끝났다고 말합니다. “우리 업계는 단기 비용, 장기 비용, 그리고 이점을 충분히 고려하지 않고 프로젝트에 좀 더 공격적으로 접근할 수 있었습니다. 이것이 가능했던 이유는 급격히 성장하는 시장에 서비스를 제공하는 것이 중요했기 때문이었습니다. 결과적으로 리소스를 확장하고 소비하기 시작했지만 정량화할 수 없는 많은 활동이 이루어졌습니다.” 그는 이제 업계가 훨씬 더 주의를 기울여야 한다고 말합니다.

Dell은 직원들이 데이터 및 AI 프로젝트를 실험하고 워크로드를 클라우드로 이전하도록 끊임없이 독려합니다. 그러나 Roese는 사용 사례에 적용할 더 엄격한 규칙이 필요합니다. 그의 메시지는 두 가지입니다. 첫째, “개념 증명에서 그 실험이 실행 가능해 보이고 이를 확장하기로 결정한 경우 활동과 관련된 단기 및 장기 비용을

모두 정량화할 수 있어야 합니다. 이 질문에 답할 수 없다면 확장을 진행해서는 안 됩니다.”

둘째, 모든 데이터 및 AI 프로젝트 제안에는 고객 가치 창출 비용 절감 등 비즈니스 영향에 대한 자세한 평가가 포함되어야 합니다. Roese는 이렇게 말합니다. “그렇지 않으면 확장을 시작한 후 ROI를 재해석해야 합니다. 이는 과거에 수행한 일과 미래의 목표를 기반으로 합니다. 그리고 AI 및 데이터 프로젝트에서는 비즈니스 혁신을 위해 노력해야 합니다.” 그는 성공을 수치화하기 위해 적용할 지표를 명확히 해야 한다고 말하면서, “측정 결과는 어느 정도 지속 가능해야 합니다.”라고 덧붙입니다.



03

현대화 및 간소화

데이터 인프라와 AI 시스템은 점점 더 얽혀가고 있는데, 주된 이유는 데이터 수집, 처리, 저장 및 분석에 대한 AI 모델의 요구 사항이 엄청나게 많기 때문입니다. 인도에 본사를 둔 금융 서비스 회사인 Razorpay의 최고 기술 책임자이자 엔지니어링 책임자인 Murali Brahmadesam은 약 5년 전 운영에 AI를 사용하기 시작한 이후 처리되는 결제량이 엄청나게 늘어났다고 말하면서, “데이터 및 처리 관련 요구 사항의 규모가 커지면서 비즈니스가 2년마다 10배씩 성장했습니다.”라고 덧붙입니다.

Brahmadesam은 이렇게 말합니다. “5년 전에는 효과가 있었던 데이터 인프라가 지금은 효과가 없습니다. 예를 들어 규모가 작을 때는 운영 데이터베이스에서 데이터를 제공할 수 있습니다. 그러나 이 정도의 규모에 도달한 후에는 고객의 요구 사항을 충족하는 동시에 비용 효율적인 방식으로 데이터를 보다 효과적으로 저장할 수 있는 성능 중심의 아키텍처를 구축해야 했습니다.” 이러한 유형의 성장은 생성형 AI로 입증된 기능을

비롯한 AI 기능의 발전과 함께 레거시 데이터 아키텍처를 현대화해야 할 필요성을 높였습니다. 또 다른 요구 사항은 아키텍처를 간소화하고, 서로 다른 데이터와 AI 플랫폼을 통합하고, 기존의 데이터 및 AI 플랫폼 전반을 포괄하도록 거버넌스를 통합하는 것입니다.

현대화

생성형 AI와 이를 지원하는 LLM은 데이터 시스템에 대한 워크로드 수요를 크게 증가시킵니다. 또한 이러한 시스템이 처리해야 하는 작업을 점점 더 복잡하게 만듭니다. Roese는 이렇게 말합니다. “생성형 AI가 데이터 아키텍처에 미치는 영향은 엄청납니다. 여기에는 3가지 과제가 수반되는데, 대부분 비정형인 엄청난 양의 데이터를 모델에 제공하여 모델을 훈련시키고, AI 소비에 도움이 되는 방식으로 데이터를 장기간 저장하며, 모델에 적절한 보안을 제공하는 것입니다.”

생성형 AI에 대한 CIO 관점에 대한 2023년 보고서인 **대규모**

“5년 전에는 효과가 있었던 데이터 인프라가 지금은 효과가 없습니다.”

Murali Brahmadesam, 최고 기술 책임자 겸 엔지니어링 책임자, Razorpay

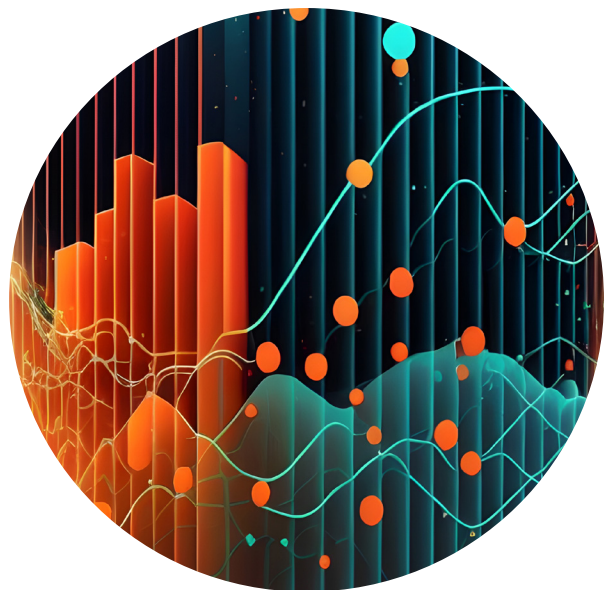
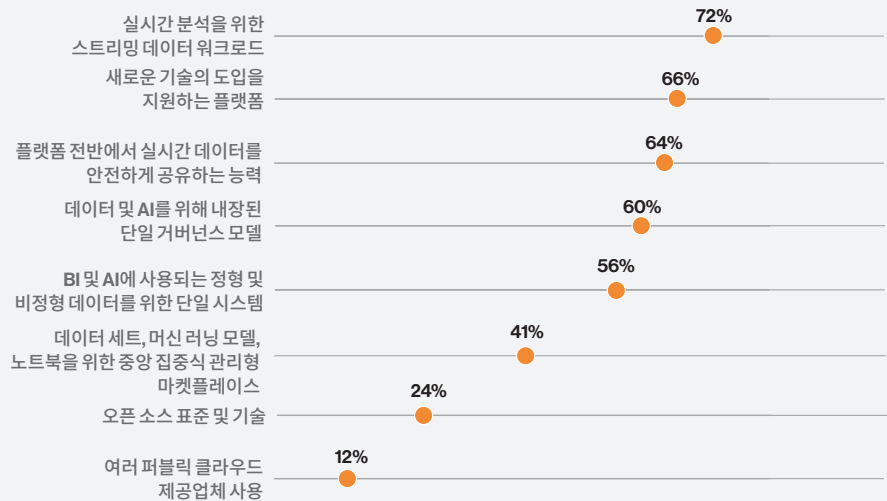


그림 3: 현대화에 필요한 인프라 속성

향후 2년 동안 조직이 전반적인 기술 목표를 달성하는 데 다음 사항이 얼마나 중요할까요?
(“매우 중요”하다고 응답한 비율(%))



출처: MIT Technology Review Insights 설문조사, 2023

가속화에서는 LLM을 지원하는 조직에 유연하고 확장 가능하며 효율적인 데이터 인프라가 필요한 이유를 집중 조명했습니다.² 이 연구를 위해 설문조사에 참여한 기술 리더들(대부분의 조직이 이미 생성형 AI를 사용하고 있음)은 데이터 인프라에 대한 요구 사항을 명확히 밝혔습니다(그림 3 참조). 이들의 우선순위는 많은 AI 애플리케이션의 기본 기능인 실시간 분석을 위한 스트리밍 데이터 워크로드(72%), 새로운 기술을 도입하는 능력(66%), 플랫폼 전반에서 실시간 데이터를 안전하게 공유하는 능력(64%), 데이터와 AI를 위한 단일 거버넌스 모델(60%)입니다.

설문조사에 참여한 기술 리더 중 3/4는 이러한 기능을 제공하기 위해 데이터 인프라를 현대화하는 과정에서 레이크하우스 아키텍처를 도입했습니다. 그렇지 않은 1/4 중 90%는 향후 3년 내에 이 아키텍처를 도입할 것으로 예상합니다(그림 4 참조). 레이크하우스는 이전에 많이 사용되었던 2가지 아키텍처 접근 방식인 데이터 웨어하우스와 데이터 레이크의 기능을 결합합니다. Brahmadേശam은 “우리가 선택한 레이크하우스 아키텍처는 비용과 성능의 적절한 균형을 유지하면서 책임 있게 확장하는 데 도움이 됩니다.”

Bhakta에 따르면 레이크하우스 아키텍처는 Condé Nast에서 데이터를 사용하여 수행하는 모든 작업의 중심에 있으며, “이를 통해 모든 데이터를 한 곳으로 모아 아키텍처를 뒷받침하는 데이터를 깔끔하게 유지할 수 있게 되었습니다.” Bhakta는 그의 팀이 레이크하우스를 통해 브론즈, 실버, 골드 데이터 계층을 갖춘 단일 데이터 플랫폼을 구축했다고 말하면서, “골드는 소비할 준비가 되었고, 실버는 소비할 준비가 되었으나, 거기에 더해 자체 시각화를 수행할 수 있으며, 브론즈는 자체적으로 정교한 분석을 실행할 수 있는 원시 데이터입니다.”라고 덧붙입니다. TSA의 현대화된 데이터 아키텍처를 통해 CIO는

수집하여 저장된 데이터를 더욱 명확하게 볼 수 있습니다. Oshinnaiye에 따르면, “이제 법률 시행 관련 데이터, 항공사 보안 검색대 데이터, 철도 데이터, 검사 데이터를 관리할 수 있습니다. 그리고 동시에 이 모든 데이터를 개별적으로 그리고 함께 상관관계를 확인하고 이해할 수 있습니다.”

간소화

오늘날 데이터 사일로는 데이터 및 기술 경영진의 골칫거리이며, 조직 내 여러 부분이 수집하거나 생성한 데이터를 저장 및 보호하기 위해 취한 서로 다른 접근 방식과 이를 관리하기 위해 IT가 획득하거나 구축한 다양한 플랫폼의 유산입니다. 내부적으로 그리고 벤더 클라우드에서 실행되는 여러 데이터, 분석 및 AI 시스템으로 인해 복잡성이 가중되었습니다. 그 결과 수많은 플랫폼, 방대한 양의 데이터 중복, 그리고 별도의 거버넌스 모델이 존재하는 경우가 많아졌습니다.

설문조사에 참여한 대부분의 조직은 10개 미만의 데이터 및 AI 시스템을 사용하고 있지만 시스템의 확산은 대규모 기업에서 가장 광범위하게 이루어졌습니다(그림 5 참조). 연간 수익이 100억 달러가 넘는 조직 중 81%는 이러한 시스템을 10개 이상 보유하고 있으며 28%는 20개 이상을 사용하고 있습니다. 인터뷰 대상 조직은 사용하는 플랫폼 수를 통합하고 전사적으로 데이터를 원활하게 연결하는 것을 목표로 합니다. Zutshi는 이러한 유형의 간소화에는 많은 이점이 있다고 말하면서, “여러 정책과 거버넌스 엔진, 여러 웨어하우스, 별도의 레이크하우스를 갖출 필요는 없습니다. 통합된 정책과 거버넌스 엔진을 갖춘 하나의 레이크하우스 플랫폼만 있으면 됩니다. 각 사용자 유형을 지원하기 위해 지속적으로 데이터를 주고 받을 필요는 없습니다. 대신 다양한 사용자 유형이 하나의 데이터 복사본을 사용하여 작업을 수행하게 됩니다.”라고 덧붙입니다.

그림 4: 레이크하우스 데이터 아키텍처의 도입

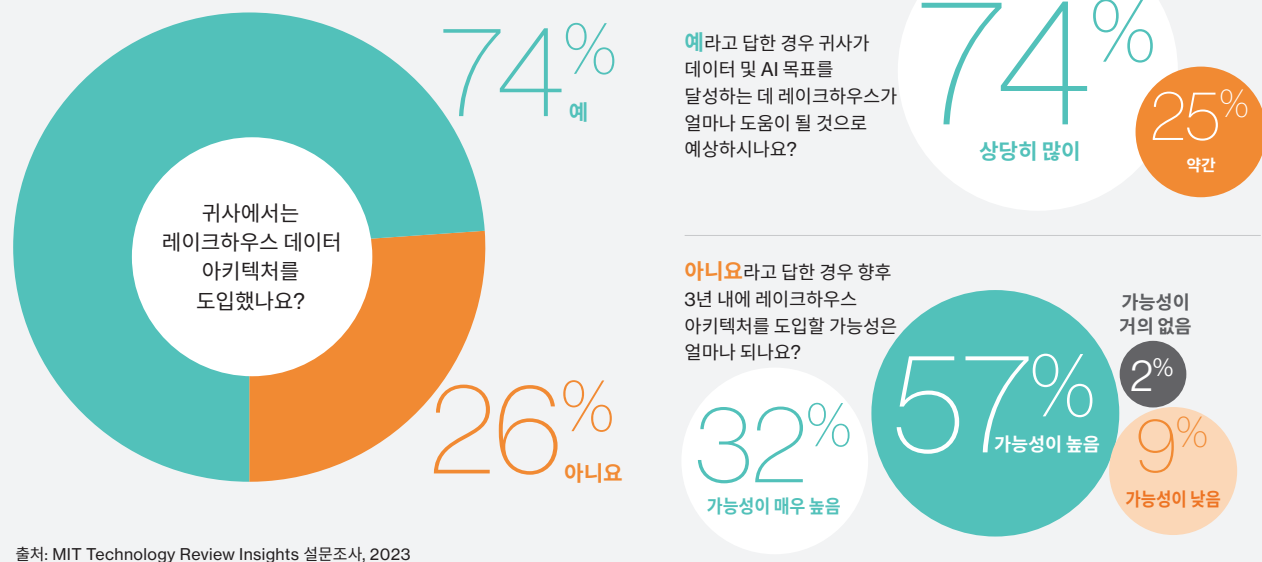
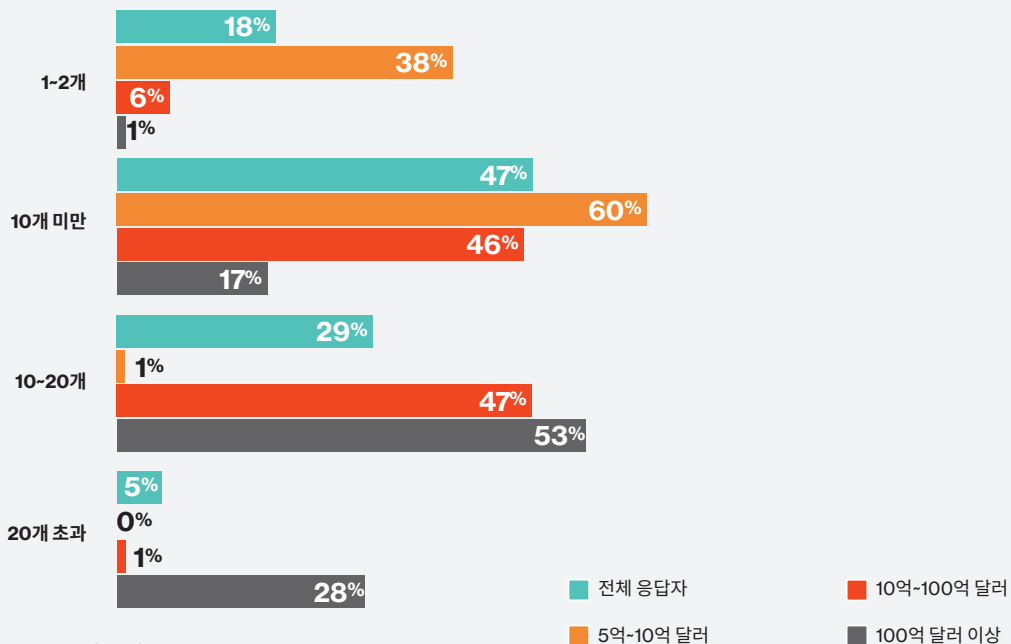


그림 5: 데이터 및 AI 시스템의 확산

현재 조직에서 사용하는 데이터, AI 및 ML 시스템은 몇 개인가요? (매출 기준)



“레이크하우스를 통해 모든 데이터를 한 곳으로 모아 아키텍처를 뒷받침하는 데이터를 깔끔하게 유지할 수 있게 되었습니다.”

Sanjay Bhakta, 최고 제품 및 기술 책임자, Condé Nast

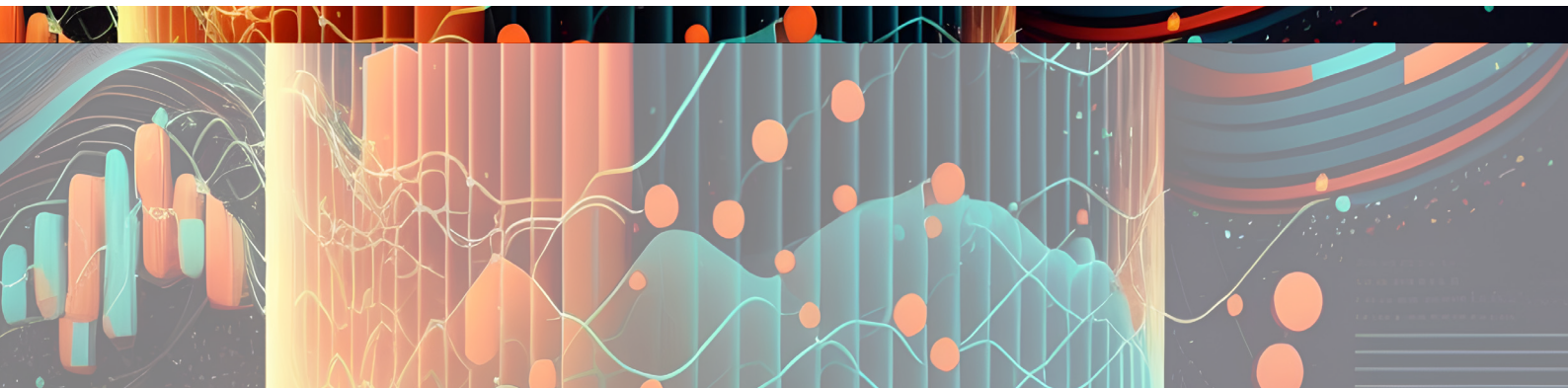
인터뷰 대상 조직은 간소화를 위해 다양한 접근 방식을 취하고 있습니다. Hall Lefevre는 스타벅스는 클라우드를 중심으로 도메인별 데이터 허브를 구축하여 조직 전반의 데이터를 중앙 집중화하고 있다고 말하면서, “이 접근 방식은 중복성을 줄이는 퍼즐의 핵심 조각이며, AI 사용 사례를 위한 강력한 기반을 제공합니다.”라고 덧붙입니다.

GM의 데이터 및 분석 팀의 경우 인프라를 간소화하고 중복을 방지하기 위해 재사용 가능한 기술에 초점을 맞추고 있습니다. Francis는 이전에 근무했던 기업에서 특정 애플리케이션을 지원하는 특수 목적 플랫폼의 확산을 목격했습니다. 그는 “AI를 배포할 때 기술 스택을 복제하거나 다른 라이선스로 맞춤형 기술을 구축하는 것을 피하기 위해 재사용 가능한 프레임워크와 기술을 고려해야 합니다.”라고 말하면서, “저희 데이터 팀은 IT와 협력하여 재사용 가능한 프레임워크를 더 많이 구축하고 있습니다.”라고 덧붙입니다.

조직에는 혁신을 위한 공간이 필요하지만, 이러한 혁신은 메인스트림 데이터 시스템 범위 밖에서 더 쉽게 이루어지는 경우가 많습니다. Regeneron Genetics Center의 최고 데이터 책임자인 Jeffrey Reid는 데이터의 중앙 집중화와 완전한 상호 운용성은 분명 바람직하지만 완전한 통합과 자유로운 혁신 사이에서 균형을 맞추는 것은 어려울 수 있다고 말하면서, “우리는 데이터 관리(데이터 통합)에 관한 일부 기능과 새로운 탐색이 많은 기타 기능을 통해 이러한 균형을 이루려고 노력합니다. 이러한 탐색으로 성공을 거두면 데이터를 한 곳으로 모드는 데 도움이 됩니다.”라고 덧붙입니다.

“AI를 배포할 때 기술 스택을 복제하거나 다른 라이선스로 맞춤형 기술을 구축하는 것을 피하기 위해 재사용 가능한 프레임워크와 기술을 고려해야 합니다.”

Jon Francis, 최고 데이터 및 분석 책임자, General Motors



04 생성형 AI 가속기

대 규모 가속화 보고서에 따르면 생성형 AI가 업계의 판도를 바꿀 수 있는 잠재력을 가지고 있는 것은 분명합니다.³ ChatGPT 및 이와 유사한 챗봇은 AI의 가능성을 일반 대중과 기술자 모두의 일상적인 대화 주제로까지 끌어올렸습니다.

Zutshi는 이렇게 설명합니다. “IT 또는 SQL 분석가가 답변할 때까지 기다리는 대신 SQL로 변환되는 영어와 유사한 질문을 하고 그에 대한 응답을 인사이트로 변환할 수 있습니다. 기업에서 자체적으로 이러한 작업을 수행할 때 생산성이 크게 향상되는 것을 보게 될 것입니다. 이 부분에 대한 관심이 고조되어 있으므로 CIO로서 활용하고 싶을 것입니다.”

마운팅 도입 과제에 주력

설문조사 대상 조직 중 약 1/4(26%)이 현재 생성형 AI 시스템에 투자하고 있고(그림 6 참조), 62%의 조직은 실험 단계에 있습니다. 설문조사에 참여한 매출이 100억 달러 이상인 조직은 그 유용성에 대해 의심의 여지가 거의 없는 것으로 보입니다. 70%는 이미 생성형 AI에 투자하고 도입하고 있는 반면 매출이 5억~10억 달러인 조직은 3%에 불과합니다.

생성형 AI의 물결은 실리콘 밸리에만 국한된 현상이 아니라 세계적인 현상입니다. 전 세계의 기업들은 경쟁 우위를 확보하기 위해 기술에 투자하고 이를 실험하고 있습니다. 싱가포르와 일본의 설문조사 응답자들은 생성형 AI를 도입하기 위해 가장 빠르게 움직이고 있고, 엄박한 EU의 규제에도 불구하고 유럽 국가들도 이 기술에 관심을 기울이고 있습니다(그림 7 참조). 개인화, 공급망 최적화 및 품질 관리는 가치를 제공할 가능성이 높은 것으로 가장 자주 언급되는 생성형 AI의 사용 사례입니다(그림 8 참조). 그러나 사용 사례의 상대적 중요성은 산업마다 크게 다릅니다. 예를 들어, 실시간 데이터 분석은 정부 기관에서 가장 많이 사용되는 사용 사례이며 금융 서비스 및

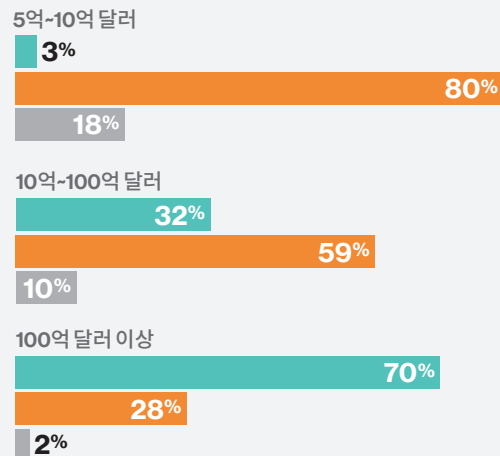
에너지 공급업체에서도 중요도가 높습니다. 제조, 리테일, 미디어 및 엔터테인먼트, 통신 업계의 경영진은 생성형 AI가 자동화와 효율성 측면에서 상당한 가치를 제공할 것으로 기대합니다. 그리고 금융 서비스와 에너지 공급업체는 이러한 모델이 제공할 수 있는 위험 관리 이점을 특히 높게 평가하고 있습니다.

인터뷰에 참여한 대부분의 기술 리더들은 현재 생성형 AI 애플리케이션을 개발 중이거나 앞으로 그렇게 할 가능성이 높습니다. 아직 결단을 내리지 않은 기술 리더들은 LLM에 관해 아직 해결되지 않은 의문점을 가지고 있는 경우가 있습니다. Oshinnaiye는 이렇게 말합니다. “LLM은 분명 우리의 미래에 영향을 미칠 것입니다. 고객 참여라는 관점에서 앞으로 일상이 되는 기술을 활용할 수 있다는 것은 매우

그림 6: 생성형 AI 도입 증가

귀사는 현재 생성형 AI를 어떻게 사용하고 있나요? (매출 기준)

- 투자 및 도입 단계
- 실험 단계
- 아무 조치도 취하지 않음



출처: MIT Technology Review Insights 설문조사, 2023

중요합니다.” 하지만 그는 기존 LLM에는 현재 이벤트에 대한 무지와 같은 한계가 있다고 지적합니다. “예를 들어 ChatGPT는 1년 이상 뒤쳐질 수 있습니다. 점점 나아지고 있지만 항상 최신 정보를 얻을 수 있는 것은 아닙니다. 그렇다면 데이터 소스와 현재 정보를 LLM에서 어떻게 활용할 수 있을까요?”

Zutshi는 특정 모델에 대한 종속은 본질적으로 제한적이라고 덧붙이면서, “모델이 계속 반복되기 때문에 한 모델에 얽매는 것을 원하지 않습니다. 효율성 향상에 따른 이점을 활용하려면 시간 경과에 따라 모델을 전환할 수 있어야 합니다.”라고 말합니다. 특수 목적의 모델과 인프라는 이 분야에서 중요한 옵션이 될 수 있으며, 이를 통해 조직은 사용된 데이터를 정리 및 조사하고 자체 독점 데이터를 제공하며 다양한 모델의 장점을 활용할 수 있습니다. Zutshi는 “Databricks는 이러한 모델에 다양한 옵션을 제공하는 게이트웨이인 추상화 계층을 구축했습니다. 따라서 작업에 맞게 올바른 모델을 선택할 수 있습니다.”라고 말합니다.

생성형 AI는 또한 거버넌스와 위험과 관련된 새로운 과제를 제기합니다. Roese는 이렇게 설명합니다. “이전 AI 시대에는 대부분의 프로젝트가 주로 도메인이나 프로세스별로 진행되었습니다. 하지만 현대에 볼 수 있는 것과 같이 일반화된 대규모 AI 프로젝트가 아니었습니다. 이러한 새로운 시대에 접어들면서 잘못된 위험이 비약적으로 증가했습니다.”

Hall Lefevre는 종합적인 거버넌스 모델이 핵심이라고 생각합니다. 그녀는 이렇게 설명합니다. “AI의 효율성은 데이터 품질과 데이터 관리 방식에 크게 좌우되기 때문에 데이터와 AI 거버넌스는 전체적으로 일관성 있게 동기화되어야 합니다.

“IT 또는 SQL 분석가가 답변할 때까지 기다리는 대신 SQL로 변환되는 영어와 유사한 질문을 하고 그에 대한 응답을 인사이트로 변환할 수 있습니다. 기업에서 자체적으로 이러한 작업을 수행할 때 생산성이 크게 향상되는 것을 보게 될 것입니다.”

Naveen Zutshi, 최고 정보 책임자,
Databricks

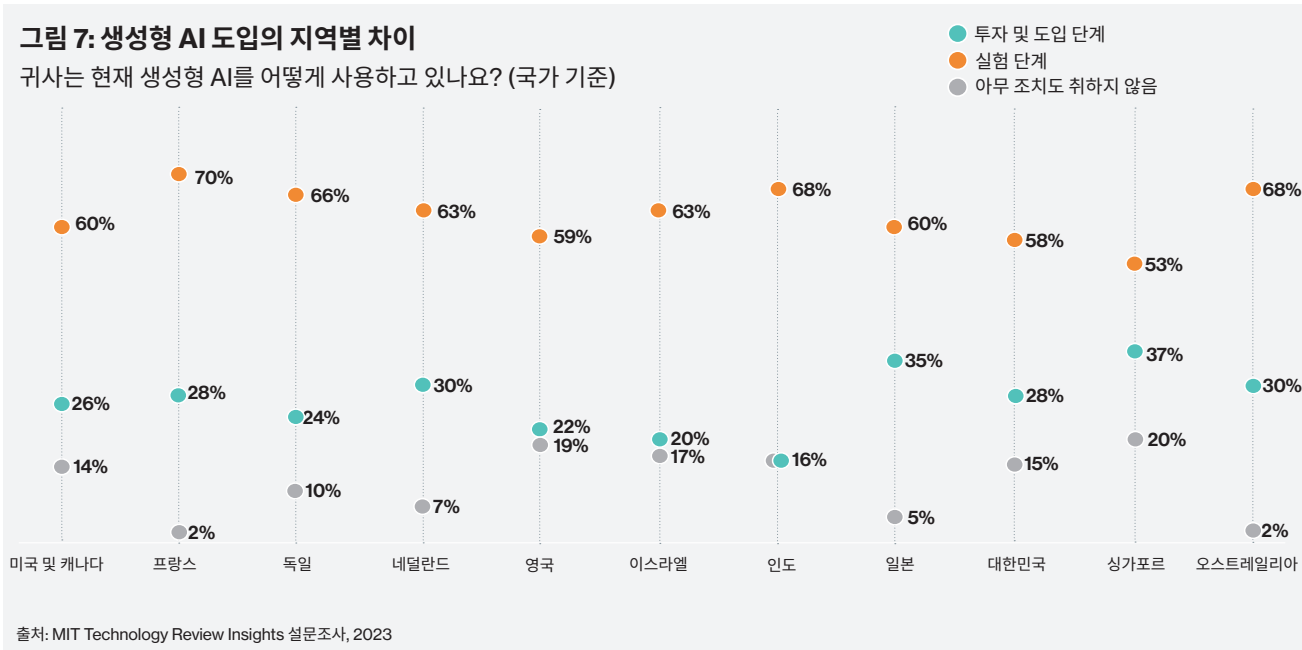


그림 8: 산업별 최고의 생성형 AI 사용 사례

향후 2년 동안 귀사의 조직에 가장 큰 가치를 가져다 줄 생성형 AI 사용 사례는 무엇일까요?

설문조사 응답자들에게 상위 3가지 옵션을 선택하도록 요청했습니다.

	전체 응답자	금융 서비스	정부 및 공공 부문	의료 및 생명 과학	리테일 및 CPG	제조업	미디어 및 엔터테인먼트	에너지	통신
개인화 및 고객 경험	1	1	3(공동)	1	2		1	3(공동)	
공급망 최적화	2			2	1	1		1	
품질 관리	3			3		2		2	1
실시간 데이터 분석 및 인사이트		2	1					3(공동)	2
자동화 및 효율성			2		3	3	2		3
혁신적인 제품 및 서비스							3		
예측적 유지관리			3(공동)						
위험 관리		3						3(공동)	

출처: MIT Technology Review Insights 설문조사, 2023

“결국 중요한 것은 머신러닝 모델에 제공되는 서비스의 데이터 품질입니다.”

Jack Berkowitz, 최고 데이터 책임자, ADP

거버넌스 모델에 대해 전체 조직을 교육하고 조율하는 것도 필수적입니다. 기업 전체에서 AI 사용이 폭발적으로 증가함에 따라 모두가 거버넌스 정책을 완벽하게 이해해야 합니다.”

Reid는 거버넌스가 통제와 활성화 사이에서 균형을 잡아야 한다고 지적합니다. “누군가 출근길에 아이디어가 떠올랐을 때 이를 테스트하기까지 일주일의 기다릴 필요가 없도록 데이터를 설정해야 합니다. 이에 우리는 보안과 개인정보를 보호하면서 민첩한 실험을 지원하기 위해 노력합니다.”

LLM은 훈련 시 엄청난 양의 데이터를 이용하며 적절한 제어가 이루어지지 않으면 민감한 데이터를 노출할 수 있습니다. 급여 및 HR 시스템 제공업체 ADP의 최고 데이터 책임자인 Jack Berkowitz는 LLM이 사용하는 데이터의 출처를 입증하는 데에도 문제가 있다고 말하면서, “2019년부터 LLM을 사용해 왔지만 최신 버전에서는 서비스 데이터의 품질 보장이 더욱 복잡해졌습니다.”라고 설명합니다. 그리고 모델이 어떤 데이터를 학습했는지 불분명하면 고객에게 제공되는 조언이 정확한지 확인하기 어렵다고 덧붙입니다.

ADP: 생성형 AI의 과제를 직시하고 있음

ADP가 2019년에 내부 LLM 기능 개발을 시작했을 때 최고 데이터 책임자인 Jack Berkowitz와 그의 팀은 데이터 아키텍처 요구 사항을 재평가해야 한다는 사실을 바로 깨달았습니다. 문제는 다양한 AI 프로젝트에서 생겨나는 여러 개의 새로운 데이터 사일로였습니다. 팀은 독립적인 데이터 파이프라인을 구축하고 API를 사용하여 서로 연결했습니다. Berkowitz는 당시를 회상하면서, 그 결과 “데이터를 계속해서 복제하면서 혼란이 발생했습니다. 데이터 최신 상태, 품질, 거버넌스 관리가 감당하기 어려워졌습니다.”라고 말합니다.

이후 2년 동안 구현된 ADP의 솔루션은 프로젝트에서 사용 중인 정형 데이터와 반정형 데이터를 통합하는 것이었습니다. Berkowitz는 이렇게 설명합니다. “이러한 데이터를 통합할 수 있다면 머신 러닝 애플리케이션을 개발하는 데 도움이 될 반복 가능한 프로세스를 구축할 수 있을 것이라고 생각했습니다. 그래서 그렇게 했습니다. 이제 우리는 데이터 레이크하우스에 비정형 데이터와 반정형 데이터의 대부분을 갖게 되었습니다. 해당 워크로드에는 10페타바이트 이상의 데이터가 있는데, 미국과 캐나다의 트랜잭션 시스템에서 15초 이내에 새로 고쳐지고, 모두 해당 국가의 데이터 보호 및 개인정보 보호 규칙을 준수합니다.”

Berkowitz에 따르면 이 아키텍처 솔루션이 제공한 가장 중요한 이점 중 하나는 그의 팀이 몇 년 전에 개발하기 시작한 분류 기준과 기술의 재사용 가능성입니다. 그는 “프로젝트 팀은 더 이상 자체적으로 구축할 필요가 없습니다.”라고 말합니다. 그러나 가장 큰 이점은 데이터 정확성과 무결성에 대한 신뢰성이었습니다. Berkowitz는 이렇게 말합니다. “결국 중요한 것은 머신 러닝 모델에 제공되는 서비스의 데이터 품질입니다.”

또 다른 생성형 AI의 과제는 완성된 모델에 적절한 보안을 적용하는 것입니다. 실제로 데이터 보안과 개인정보 보호는 생성형 AI의 사용에 대한 응답자의 주요 관심사이며, 응답자 중 50%가 이에 동의했습니다. Roese는 이렇게 말합니다. “훈련된 모델은 엄청나게 가치가 있습니다. 그렇다면 랜섬웨어 공격이 발생할 경우 누구도 접근할 수 있도록 어떻게 보장할 수 있을까요? 그리고 프로덕션으로 전환할 때 해킹될 수 없는 플랫폼에서 작동하고 있나요?”

그러나 Zutshi는 AI가 보안에 미치는 영향에는 장단점이 모두 있다고 지적합니다. 그는 이렇게 설명합니다. “보안은 긍정적인 측면에서 AI로 인해 가장 많이 혼란을 겪게 될 영역이라고 생각합니다. 왜냐하면 비정상적인 행동을 탐지하고 이에 대해 훨씬 더 빨리 조치를 취하는 데 사용할 수 있는 엄청난 양의 데이터가 있기 때문입니다. 모든 회사의 보안 태세에 AI 기반 자동화가 훨씬 더 큰 영향을 미치게 될 것이라고 믿습니다.”

구축 또는 구매

조직은 자체 LLM을 내부적으로 구축할지 아니면 기술 벤더가 제공하는 독점 모델을 사용할지 결정해야 합니다. 설문조사에 참여한 경영진은 대부분 하이브리드 접근 방식을 취하고 있으며, 58%는 구축과 구매 모두를 고려하고 있다고 밝혔습니다. 13%는 오픈 소스 모델을 기반으로 자체 LLM을 구축하기로 결정했으며, 29%는 외부 모델만 사용한다고 답했습니다(그림 9 참조).

설문조사 표본 중 가장 작은 조직은 생성형 AI 구축을 위한 유일한 접근 방식으로 벤더 모델을 채택할 가능성이 가장 큰 조직의 2배입니다. 이는 내부적으로 구축할 때 비용이 높다는 인식으로 인해 많은 조직이 이 옵션을 선택하지 않는다는 것을 의미합니다. 그러나 LLM 훈련을 위한 저렴한 오픈 소스 옵션이 등장함에 따라 비용에 대한 우려는 줄어들 것입니다.

많은 조직에서 벤더 LLM을 사용하면 IP 소유권, 보안, 개인정보 보호 및 정확성을 포함한 위험이 따릅니다. 따라서 어떤 옵션을 사용할지에 대해 케이스별로 결정을 내릴 수 있는 하이브리드 접근 방식이 적합합니다. Brahmedesam은

“LLM은 분명 우리의 미래에 영향을 미칠 것입니다. 고객 참여라는 관점에서 앞으로 일상이 되는 기술을 활용할 수 있다는 것은 매우 중요합니다.”

Yemi Oshinnaiye, 최고 정보 책임자, U.S. Transportation Security Administration

이렇게 설명합니다. “공개적으로 교육을 받은 LLM은 개발자 생산성과 같은 사용 사례에 가장 적합합니다. 그러나 고객 유형 파악(KYC) 제도의 경우 모델이 분석할 문서가 공개적으로 제공되지 않기 때문에 우리의 요구 사항에 맞는 고유한 모델을 구축해야 합니다. KYC에서는 정확성이 필수적입니다.”

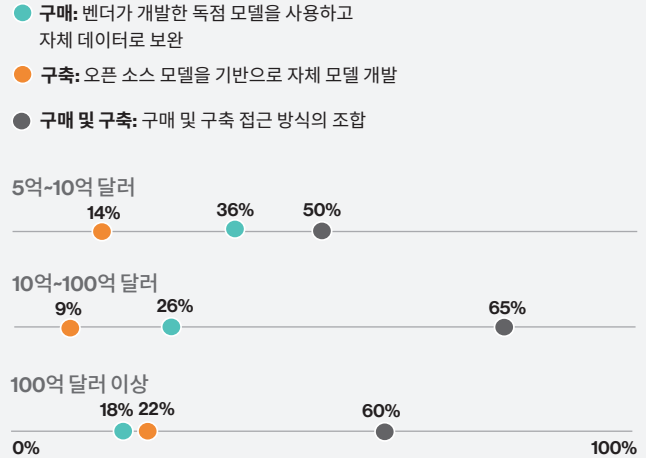
스타벅스의 경우에도 하이브리드 접근 방식이 적합합니다. Hall Lefevre는 어떤 옵션을 선택하든 보안이 가장 중요하다고 말합니다. “모델을 채우고 조정하기 위해 사용하는 모든 데이터가 매우 안전한 자체 클러스터에 있도록 하고 싶습니다. 오픈 소스 벤더를 선택한다면 IP를 유지할 수 있도록 해당 업체의 제품과 도구가 클러스터에서 안전하게 실행될 수 있는지 확인해야 합니다.”

Reid에 따르면 생명 공학 산업에서는 내부 구축이 필수인 경우가 많습니다. “과학적 발견에는 ChatGPT와 같은 공개 모델에 사용되는 토큰화 스키마에 표시된 것보다 생물학적 개념을 더 간결하게 표현해야 하기 때문에 내부 구축이 필요합니다.” 그럼에도 불구하고 그는 덜 까다로운 사용 사례에 파운데이션 모델을 사용할 수 있기를 바랍니다. 그는 이렇게 말합니다. “이미 우리가 할 수 있는 것보다 더 훌륭하게 뭔가를 해내는 파운데이션 모델이 있는데, 그것으로 시작하지 않을 이유가 있을까요?”

그림 9: 생성형 AI 개발에 대한 접근 방식

생성형 AI/LLM을 도입할 때 귀사에서 어떤 접근 방식을 취하거나 취할 가능성이 있나요?

(매출 기준)



“오픈 소스 벤더를 선택한다면 IP를 유지할 수 있도록 해당 업체의 제품과 도구가 클러스터에서 안전하게 실행될 수 있는지 확인해야 합니다.”

Deb Hall Lefevre, 영업 담당 부사장 겸 최고 기술 책임자, Starbucks

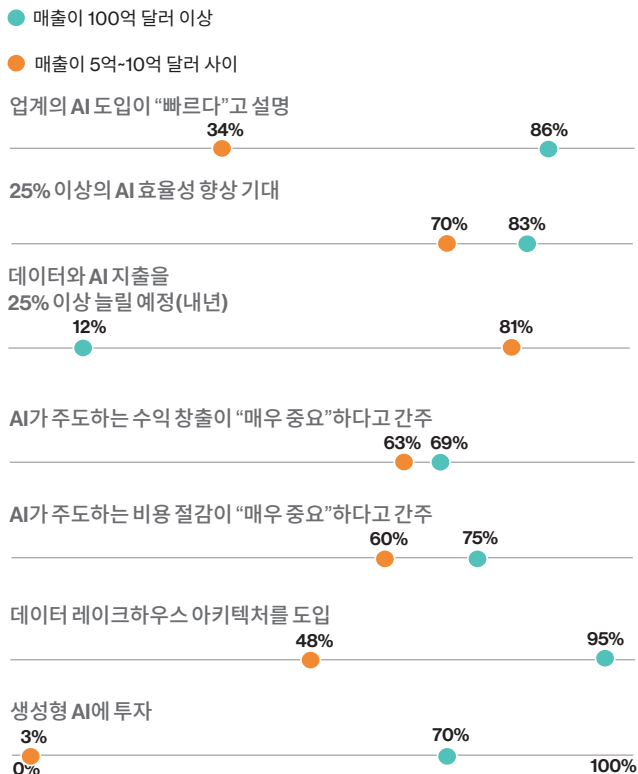
대기업을 위한 AI 확장성의 이점

설문조사 전반에 걸쳐 데이터와 AI 기능의 여러 측면에서 대규모 기업과 소규모 기업 표본 간에 극명한 차이점들이 나타났습니다. 일부 차이점은 단순히 조직이 원하는 대로 투자할 수 있는 재정적 자원으로 인해 발생할 수 있습니다. 예를 들어 대규모 기업의 70%가 생성형 AI에 투자한 반면 소규모 기업은 3%에 불과했습니다. 그리고 대규모 기업의 응답자 10명 중 8명은 내년에 데이터 및 AI에 대한 지출을 25% 이상 늘릴 것으로 예상하는 반면 소규모 조직에서는 그렇게 할 의도가 있는 비율이 훨씬 더 적습니다(12%)(그림 10 참조).

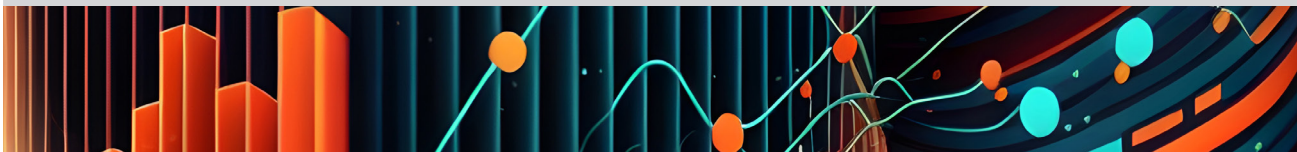
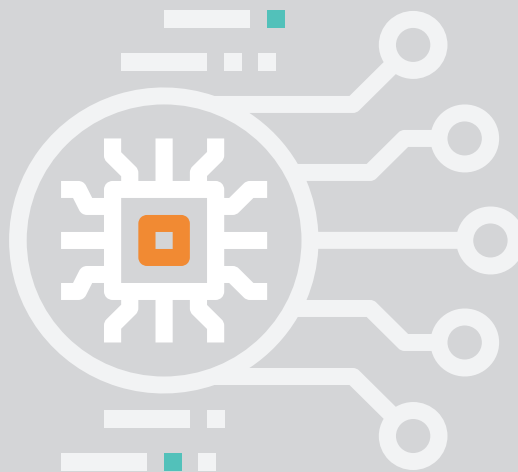
이러한 차이는 전망의 차이를 반영할 수도 있습니다. 예를 들어 대규모 조직의 기술 경영진은 업계의 AI 도입 속도에 대해 소규모 조직의 동료보다 확실히 더 낙관적입니다. 대규모 조직의 86%는 속도가 “빠르다”고 설명하는 반면, 소규모 조직의 경우 34%만 “빠르다”고 설명합니다. 이는 소규모 기업의 응답자 중 향후 2년 내에 AI가 주도하여 업계의 효율성을 크게 향상시킬 것으로 예상하는 사람이 적은 이유를 설명하는 데 도움이 될 수 있습니다. 그러나 경쟁업체가 데이터 및 AI 기능 구축에서 앞서가는 것을 경영진이 보지 못한다면 산업 전반의 도입에 대한 낮은 기대치가 투자 의도에 영향을 미칠 수도 있습니다.

이러한 차이점은 소규모 기업이 AI를 마스터하기 위한 경쟁에서 대규모 기업에 뒤처져 돌이킬 수 없을 상황이 될 위험이 있음을 시사합니다. 제한된 리소스로 인해 일부에서는 주의가 필요할 수 있지만, 비교적 저렴한 비용으로 생성형 AI 모델 및 기타 클라우드 기반 AI 리소스에 액세스할 수 있는 오픈 소스 옵션을 사용할 수 있습니다. 소규모 기업은 상대적으로 고객 경험 개선, 운영 효율성 향상, 제품 혁신을 통해 대기업과 마찬가지로 많은 이점을 얻을 수 있습니다. 그들은 AI가 제공하는 성장 기회를 그냥 지나칠 여유가 없습니다.

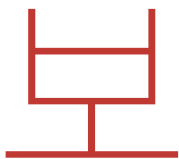
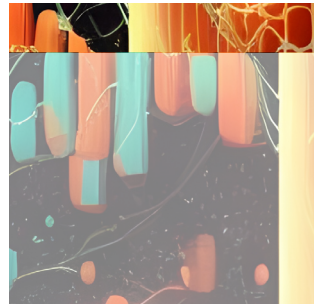
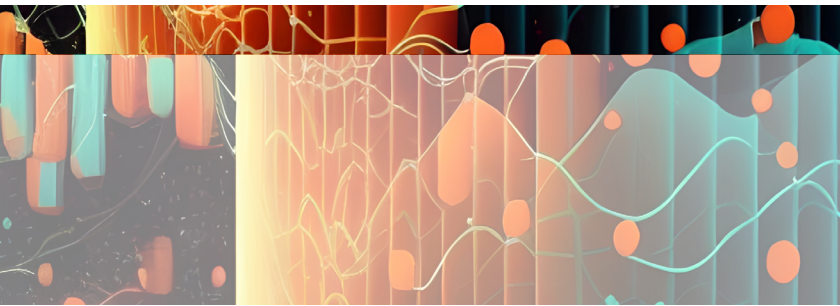
그림 10: AI 공백: 대기업과 소규모 기업의 설문조사 응답



출처: MIT Technology Review Insights 설문조사, 2023



05

시티즌(citizen)
개발자의 부상

편화라는 용어는 클라우드와 SaaS(Software-as-a-Service)가 뿌리를 내리고 모든 상업부의 직원들이 업무용 애플리케이션에 손쉽게 액세스할 수 있게 되면서 기업 기술 용어집에

포함되었습니다. 직원들이 여러 사업부 전반의 데이터에 쉽게 액세스하는 것도 중요하다는 사실이 곧 분명해졌습니다. 그 이후 업계 전반의 기술 경영진은 데이터의 보편화와 분석에서 얻은 인사이트를 현실화하기 위해 노력해 왔습니다.

이를 지원하기 위해 설문조사 응답자의 64%는 플랫폼 전반에 걸쳐 실시간 데이터를 안전하게 공유할 수 있게 해주는 데이터 아키텍처가 필요할 뿐만 아니라 그들의 기술 목표 달성에 “매우 중요”하다고 생각합니다. 그 외 41%는 데이터 세트, ML 모델 및 노트북을 위한 중앙 집중식 관리형 마켓플레이스를 갖는 것이 매우 중요하다고 생각합니다(그림 3 참조). API와 같은 데이터 공유 소프트웨어와 데이터

“생성형 AI는 AI 보편화를 위해 내딛을 첫 번째 큰 도약입니다. 우리 조직은 이에 대해 매우 낙관적입니다.”

Naveen Zutshi, 최고 정보 책임자,
Databricks

레이크하우스와 같은 아키텍처의 발전은 데이터 보편화의 촉매제로 입증되었습니다. Bhakta에 따르면 브론즈, 실버, 골드 계층으로 구성된 Condé Nast의 단일 데이터 플랫폼의 주요 목표는 “조직 전반의 데이터를 보편화”하는 것이었습니다.

ADP가 조직 내 다양한 데이터 저장소를 통합 플랫폼으로 통합하기 시작했을 때 Berkowitz는 모든 소유자의 의견을 거스르게 될 것이라는 말을 들었습니다. Berkowitz는 이렇게 설명합니다. “정반대의 일이 일어났습니다. 현재 60~70개 팀이 플랫폼을 통해 독립적으로 업무를 수행하고 데이터를 공유하며 협업하고 있습니다.”

AI 보편화

AI는 오늘날 조직 내 보편화의 새로운 목표입니다. Zutshi는 이렇게 말합니다. “생성형 AI는 AI 보편화를 위해 내딛을 첫 번째 큰 도약입니다. 우리 조직은 이에 대해 매우 낙관적입니다.” Francis도 이에 동의합니다. 그는 “6개월 전만 해도 AI가 보편화되려면 아마도 10년은 더 걸린다고 말했을 것입니다.”라면서, “하지만 지금 우리 눈앞에서 생성형 AI가 활용되고 있습니다.”라고 덧붙입니다. GM에서 그의 팀은 생성형 AI를 “조력자”로 활용하여 다른 사업부의 직원이 코드를 개발하고 애플리케이션을 구축하도록 돕는 방법을 모색하고 있습니다. 그는 “물론 선임자들이 코드를 검토해야 하지만, 그러한 역량을 갖추면 AI를 확장하고 AI 사용을 통해 혁신의 범위를 확대하는 데 도움이 될 것입니다.”라고 말합니다.

Reid는 생성형 AI 모델의 간단한 텍스트 및 음성 인터페이스는 결국 오늘날 소프트웨어 개발자가 사용하는 전문 인터페이스를 대체할 것이라고 말하면서, “보다 일상적인 인터페이스를 사용하면 코딩 기술이 없는 사람이라도 누구나 쉽게 데이터를 쿼리할 수 있습니다.”라고 덧붙입니다. Reid는 생성형 AI의 가치는 분석 그 이상이라고 말하면서, “의사 결정을 지원하기 위해 무언가를 기억하거나 이미 알고 있는 것을 검증하는 데 도움이 되는 기회를 제공합니다. 그리고 전반적으로 모든 사람을 지원하고 있습니다.”라고 덧붙입니다.

물론 AI 보편화에는 위험이 따릅니다. 이러한 위험은 법률, 규정 준수 및 보안 팀에는 익숙할 수 있지만 생성형 AI의 도입이 증가함에 따라 특히 더 크게 다가올 것입니다. Francis는 이렇게 설명합니다. “현재 조직 내 누구나 AI 애플리케이션을 구축하도록 지원할 수 있습니다. 그렇다면 우리는 모든 법률, 개인정보 보호, 보안에 따른 영향이나 브랜드 관점에서 상업적인 측면을 충분히 고려하고 있을까요?”

AI로 인한 피해로부터 조직과 소비자를 보호하는 것에 대한 경영진의 결정은 정책 입안자와 규제 기관의 영향을 받을 가능성이 높습니다. 고위 경영진은 정부가 개입하기 전에 기존 정책을 검토하고 생성형 AI 시대에 대비할 수 있는지 확인하는 것이 좋습니다.

조력자의 참여

기술을 사용하여 사람들이 혁신하고 가치를 창출하도록 역량을 강화하는 것이 보편화의 주요 목표입니다. 그러나 사람들은 기술을 보강하여 기술과 사람의 의사 결정이 밀접하게 얽히도록 합니다.

설문조사에 참여한 기술 경영진은 자신의 조직이 데이터와 AI의 인간적 측면에 충분히 집중하지 못하고 있다고 우려하고 있습니다. 회사의 데이터 전략을 어떻게 개선해야 하는지 묻는 질문에 거버넌스 및 데이터 처리 속도와 같은 중요한 영역보다 인재와 인력에 대한 투자가 가장 높은 선택을 받았습니다(응답자의 39%가 언급). 훨씬 더 많은 비율(72%)이 인재를 유치하고 유지하는 데 도움이 될 혁신을 장려하는 것이 “매우 중요”할 것이라고 답했습니다.

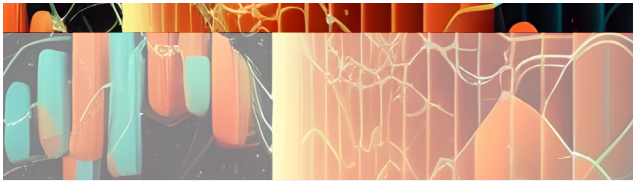


그림 11: 데이터 및 AI 플랫폼 관련 문제
귀사가 현재 갖고 있는 데이터 및 AI 플랫폼의 가장 큰 문제점은 무엇인가요?

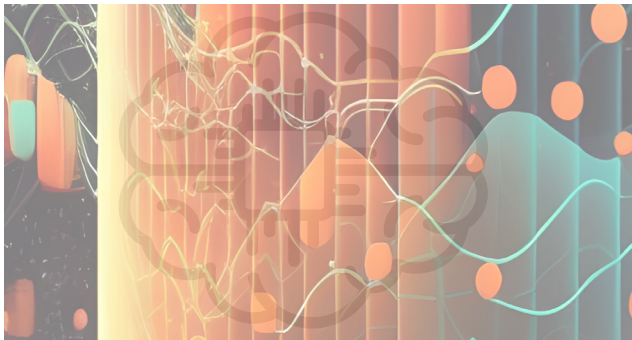
플랫폼 사용에 관한 인력 교육 또는 인력의 역량 강화	40%
부적절한 거버넌스 프레임워크	26%
사일로화된 레거시 시스템	25%
부적절한 보안 프레임워크	25%
너무 많은 이기종 시스템	18%

출처: MIT Technology Review Insights 설문조사, 2023

현재 데이터 및 AI 플랫폼에 대해 가장 우려되는 것이 무엇인지 묻는 질문에 경영진은 기술의 약점보다 인재의 병목 현상을 더 많이 선택했습니다(그림 11 참조). 응답자들은 이러한 플랫폼에서 경험하고 있는 가장 큰 “고충”으로 직원 교육 및 기술 향상의 필요성(40%)을 꼽았고 그 다음으로는 부적절한 거버넌스(26%), 사일로(25%), 부적절한 보안 프레임워크(25%) 등의 순으로 기술 관련 문제를 선택했습니다.

인력 문제는 기술 리더에게 생성형 AI 조력자 역할의 필요성을 피력하는 데 도움이 됩니다. 조직이 데이터 사이언스 분야의 인재 공백을 채우기 위해 고군분투하는 가운데 조력자 부조종사 역할이 해결책이 될 수 있다고 믿습니다. “기술을 활용하여 비즈니스 전반의 사람들이 애플리케이션을 구축하는 데 도움을 줄 수 있는 곳이라면 어디든 추가할 수 있을 것이라고 생각합니다.”

Reid는 조직 내 모든 사람이 AI 사용 방법을 배워야 한다고 굳게 믿습니다. “아주 어린 나이의 과학자가 배우는 것은 현장에서 무슨 일이 일어나고 있는지 끊임없이 인지하는 것입니다. 항상 호기심을 가져야 합니다. 그리고 모두가 AI를 포함하여 새로운 것을 배우고 훈련하기 위해 끊임없이 도전해야 합니다. 결국 이것이 과학이 발전한 방식입니다.”



“6개월 전만 해도 AI가
보편화되려면 아마도 10년은
더 걸린다고 말했을 것입니다.
하지만 지금 우리 눈앞에서
생성형 AI가 활용되고
있습니다.”

Jon Francis, 최고 데이터 및 분석 책임자,
General Motors

Databricks: 비즈니스 전반에 걸쳐 생성형 AI 운영

전사적으로 생성형 AI를 사용하려면 CIO는 비즈니스를 혁신하고, 전략적 사용 사례를 개발하고, 이러한 도구가 직원들의 진정한 조력자 역할을 하게 만드는 LLM을 운영하는 방법을 찾아야 합니다. Databricks의 CIO인 Naveen Zutshi는 가능한 한 많은 사용 사례를 만들어야 한다고 주장합니다. “생성형 AI에 대한 제 접근 방식은 아이디어에 대한 포트폴리오 접근 방식을 취하는 것입니다. 수백 가지 아이디어가 있다고 가정해 보겠습니다. 아마도 그 중 10~20개가 실제로 실현되겠죠. 여러분은 회사에서 실패하지 않는 아이디어를 가능한 많이 추진하고 싶을 것입니다.”

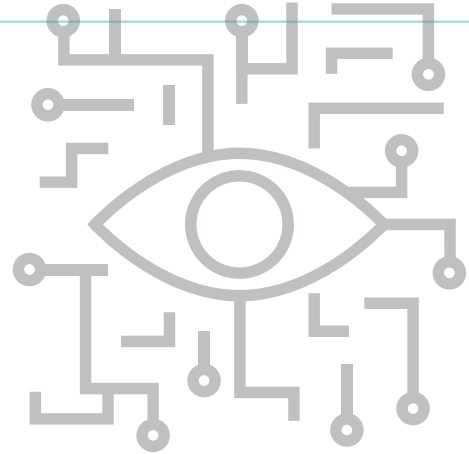
생성형 AI와 같은 기술이 대중의 상상력을 독보적으로 사로잡았을 때 많은 직원들은 자신의 업무에서 문제를 해결할 수 있는 방법을 조사하는 데 열중했습니다. Zutshi는 이렇게 말합니다. “이전에는 여러분이 회사에 ‘이런 아이디어가 있는데 우리가 이 아이디어를 추진할 수 있을까요?’라고 묻곤 했습니다. 하지만 지금은 회사에서 ‘생성형 AI와 관련해 우리가 무엇을 하고 있나요?’라고 묻고 있습니다. 비즈니스에서 끌어낼 무언가가 있고 그 열정을 활용해야 합니다.”

Databricks는 기술 팀에 국한되지 않고 영업, HR, 법무, 재무 등의 이해관계자가 참여할 수 있는 생성형 AI 해커톤을 개발하여 이러한 인기 모멘텀을 활용했습니다. 팀이 실제 비즈니스 문제와 직접적으로 관련 있는 아이디어를 클라우드소싱할 수 있게 된 것 외에도, 이러한 폭넓은 참여는 비즈니스 가치 검증 및 위험 관리와 같은 주요 비기술적 제품 요소가 처음부터 개발된 도구에 내장되어 있었음을 의미했습니다. 무엇보다도 이를 통해 비즈니스의 변화 관리를 위해 사전에 구축된 생성형 AI 챔피언이 탄생했다는 것입니다.

첫 번째 해커톤의 결과로 참가자들은 회의 요약, 계약 조항 검색, 이력서, 직무 분석부터 문서 생성 및 문제 해결에 이르기까지 다양한 사용 사례에서 효과를 발휘하는 15개의 생성형 AI 기반 소프트웨어 도구를 제작했습니다. 이 중 3개는 비즈니스 전반에 완전히 반영될 예정이고, 이러한 해커톤은 6개월 간격으로 계속 진행될 예정입니다. Zutshi는 이러한 접근 방식이 직원의 기술 역량을 강화하는 데 도움이 되고 비즈니스 내에서 AI를 확장할 수 있는 문화를 조성하여 결국 “기업 전체의 교육 의식을 고취”하는 데 도움이 될 것이라고 낙관하고 있습니다.

06

결론



AI가 경제 전반으로 훨씬 더 광범위하게 확산된다면 생성형 AI는 기술 변곡점이 될 것입니다. 전문가들은 이것이 생산성을 한 차원 더 끌어올리고 산업 전반에 걸쳐 수조 달러의 새로운 가치를 추가할 것이라고 전망합니다.⁴ 이러한 효과가 곧 실현되겠지만 생성형 AI는 조직의 데이터 및 기술 인프라와 기술 경영진이 현대화하기 위해 내리는 투자 결정에 영향을 미치고 있습니다.

AI 기능의 발전에 따른 인프라 수요로 인해 이러한 의사 결정에 관련된 이해 관계가 늘어나고 있습니다. Roese가 말한 것처럼 “잘못될 위험이 비약적으로 증가합니다.” 반면에, 생성형 AI가 가치 창출에 미치는 영향에 대한 전망 중 일부만 정확하다고 입증해도 이를 올바르게 수행할 경우 얻을 수 있는 이점은 엄청날 것입니다.

그러려면 조직의 AI 기능을 지원하는 데이터 및 기타 인프라에 대한 몇 가지 요소가 필수적입니다. 이 연구의 결과를 통해 3가지 요소가 명확한 윤곽을 드러냈습니다. 현대적이고 미래 지향적인 데이터 인프라는 다음을 보장해야 합니다.

복잡성 속의 단순성. 조직이 이러한 기능을 배포한 이후 등장한 여러 AI 및 ML 시스템과 이를 지원하는 데이터 플랫폼(예: 레이크하우스)을 통합하면 모델과 모델에 데이터를 제공하는 데이터 소스 간의 일치성이 더욱 높아질 것입니다. 그리고 사용자 인터페이스의 간소화는 AI 보편화를 촉진할 것입니다.

통제력 강화. 조직이 수년 동안 실현을 위해 노력해 온 데이터의 보편화와 함께 AI의 보편화는 새로운 가치를 창출하는 혁신을 실현하는 데 필요합니다. 그러나 모든 형태의 보편화에는 혼란을 방지하는 안전 장치가 필요하며, 이것도 예외는 아닙니다. 특히 생성형 AI의 사용 용이성은 강력하고 통합된 거버넌스의 케이스를 강화합니다.

신뢰가 바탕이 된 공유. CIO는 팀이 플랫폼 전반에서 실시간 데이터를 안전하게 공유할 수 있는 기능의 중요성을 강조합니다. 이러한 기능은 기업 내에서 분석과 AI 모델을 결합하고 조직의 공급업체와 파트너 전반을 포괄하도록 더욱 광범위하게 확장되는 데이터 에코시스템의 기반입니다. 다양한 산업에 대해 이 연구가 갖는 의미를 평가하는 후속 보고서에서 이러한 산업 데이터 에코시스템의 출현에 대해서도 자세히 알아볼 것입니다.

생성형 AI가 가치 창출에 미치는 영향에 대한 전망 중 일부만 정확하다고 입증해도 이를 올바르게 수행할 경우 얻을 수 있는 이점은 엄청날 것입니다.

MIT Technology Review Insights

MIT 테크놀로지 리뷰 인사이트는 세계에서 가장 오래 운영된 기술 매거진, *MIT 테크놀로지 리뷰*의 맞춤 출판 사업부이며, 세계에서 가장 중요한 기술 기관의 후원을 받아 당대의 주요 기술 및 비즈니스 문제에 대한 라이브 이벤트와 연구를 진행합니다. MIT 테크놀로지 리뷰 인사이트는 미국과 해외에서 양적, 질적 연구와 분석을 수행하고 기사, 보고서, 인포그래픽, 동영상, 팟캐스트를 비롯한 다양한 콘텐츠를 공개합니다. 점점 규모를 더해가는 MIT 테크놀로지 리뷰 글로벌 인사이트 패널을 통해 전 세계의 고위 임원, 혁신가, 기업가에게 설문조사와 심층 인터뷰를 의뢰하고 있습니다.

Databricks 정보

Databricks는 데이터 및 AI 기업입니다. Comcast, Condé Nast 및 Fortune 500대 기업의 50% 이상을 포함하여 전 세계적으로 10,000개 이상의 조직이 Databricks 레이크하우스 플랫폼을 사용하여 데이터, 분석 및 AI를 통합하고 있습니다. Databricks는 샌프란시스코에 본사를 두고 세계 각지에 지사를 운영하고 있습니다. Apache Spark™, Delta Lake와 MLflow를 개발한 주역들이 모여 설립한 Databricks는 데이터 팀에서 전 세계의 어렵기로 손꼽히는 문제를 해결하도록 돕는 것을 사명으로 삼고 있습니다. 자세히 알아보려면 **X**(이전 명칭: Twitter), **LinkedIn**, **Facebook**에서 Databricks를 팔로우하세요.



각주

1. "Gartner, 2023년 전 세계 IT 지출이 4.3% 증가할 것으로 전망", Gartner, 2023년 7월 19일, <https://www.gartner.com/en/newsroom/press-releases/2023-07-19-gartner-forecasts-worldwide-it-spending-to-grow-4-percent-in-2023>.
2. "대규모 가속화: 생성형 AI에 대한 CIO 관점", MIT Technology Review Insights, 2023년 7월, <https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report>.
3. "대규모 가속화: 생성형 AI에 대한 CIO 관점", MIT Technology Review Insights, 2023년 7월, <https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report>.
4. "생성형 AI의 경제적 잠재력", McKinsey & Company, 2023년 6월 14일, <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>.

그림

모든 아트웍은 Adobe Stock 및 The Noun Project 아이콘으로 제작했습니다.

이 정보의 정확성을 확인하기 위해 최선의 노력을 다하였으나, MIT 테크놀로지 리뷰 인사이트에서는 이 리포트나 이 리포트에 제시된 다른 정보, 의견 또는 결론을 활용하는 것에 대한 책임을 지거나 배상하지 않습니다.

© Copyright MIT Technology Review Insights, 2023. All rights reserved.



MIT Technology Review Insights

www.technologyreview.com

insights@technologyreview.com