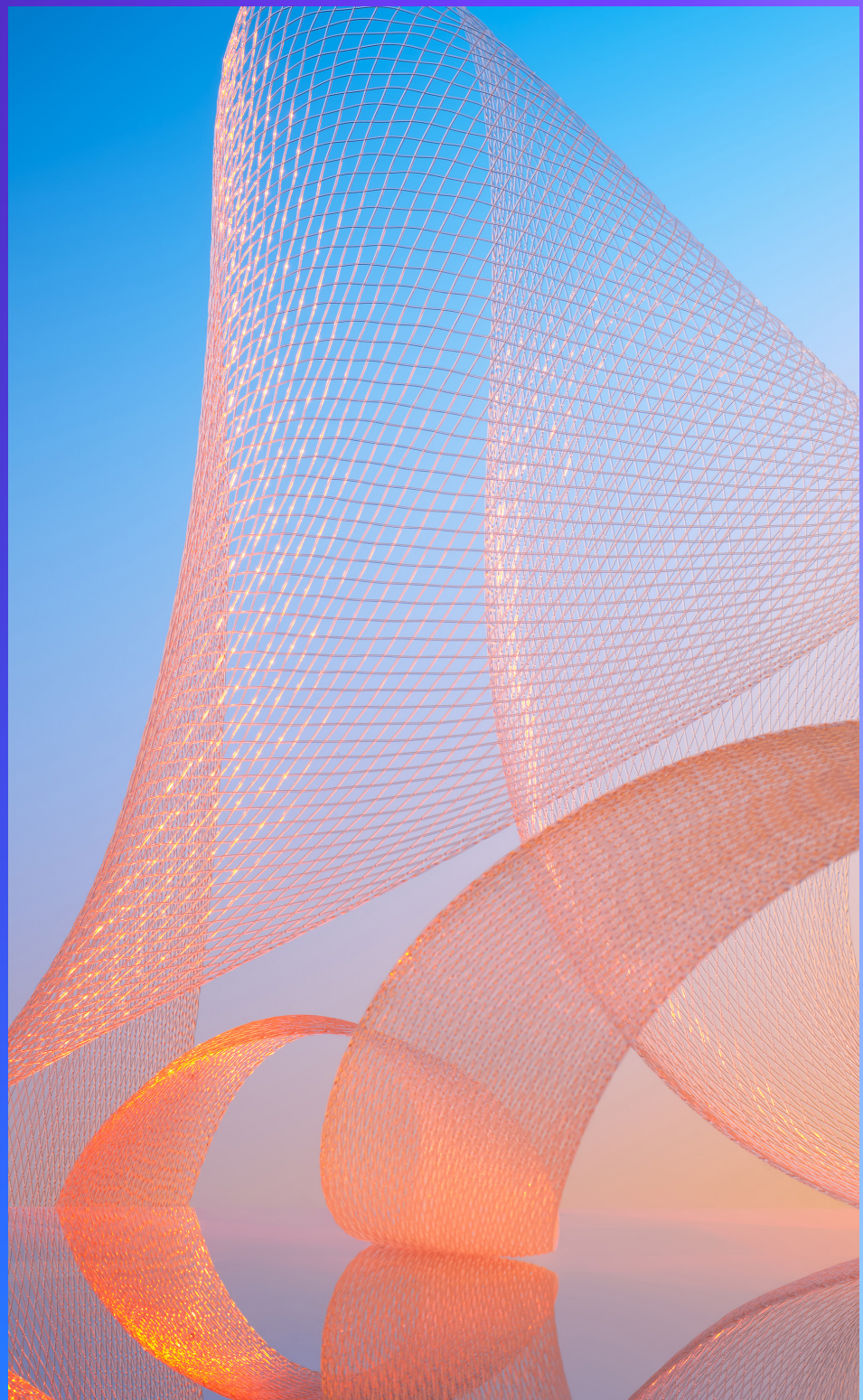
비즈니스 혁신 개요

생성형 AI로 비즈니스 가치 극대화

최대한 활용하는 생성형 AI의 잠재력



요약

소비자와 기업의 관심이 높아짐과 동시에 간편하고도 폭넓은 활용이 가능해지면서 생성형 AI(인공 지능)는 변곡점을 맞이하게 되었습니다. 오늘날 모든 조직은 생성형 AI의 영향과 그 혁신적인 가치를 활용할 방법을 평가해야 합니다.

대부분은 지금이 중요한 시점이며, 생성형 AI 탐색 전략을 신속하게 개발해야 한다는 것을 인지합니다. 그러나 아직도 다음 질문에 답하지 못하는 사람이 많습니다.

- 생성형 AI란 무엇입니까?
- 이전 세대의 AI와 생성형 AI는 어떻게 다른가요?
- 주요 비즈니스 사용 사례는 무엇인가요?
- 어떤 고객이 생성형 AI를 사용하고 있나요?
- 어떻게 시작해야 하나요?
- 위험은 무엇이며, 어떻게 완화할 수 있나요?

이 비즈니스 혁신 개요 문서는 생성형 AI의 기능, 사용 사례, 비즈니스 가치를 간략하게 안내합니다. 또한, AI 기술 및 기계 학습(ML) 도구에 대한 폭넓은 지식과 경험을 활용하는 Amazon Web Services(AWS) 분야별 전문가의 유용한 인사이트를 제공합니다.

콘텐츠 대상

이 혁신 개요 문서는 생성형 AI를 효과적으로 이해하고 활용하여 비즈니스 성과를 개선하는 방법을 배우려는 비즈니스 리더를 위해 마련되었습니다.



목차

소개: 새로운 인텔리전스의 세계 4

생성형 AI의 이해 7

생성형 AI의 비즈니스 기능 9

생성형 AI를 위한 비즈니스 고려 사항 10

생성형 AI에 대한 경영진 인사이트 12

생성형 AI의 일반적인 산업별 사용 사례 13

AWS로 생성형 AI의 성공을 이끄는 방법 16

다음 단계 20



소개

새로운 인텔리전스의 세계

소비자와 비즈니스는 정교한 콘텐츠를 만들고, 코드를 생성하고, 질문에 답하는 등 다양한 작업을 수행하는 생성형 AI의 기능에 감탄을 금치 못합니다. 간단한 자연어 프롬프트로 시작하여 대부분 몇 초 만에 작업이 완료됩니다.

지금은 소비자의 생성형 AI 활용에 대한 관심이 뜨겁지만, 사실 더 큰 기회는 비즈니스를 향해 열려 있습니다. 비즈니스는 생성형 AI를 활용하여 고객과 직원에게 멋진 경험을 선사할 수 있습니다. 생성형 AI의 진정한 힘은 검색 엔진이나 챗봇을 넘어 기업과 조직 운영 방식의 모든 측면을 변화시킬 것입니다. Gartner는 2025년까지 대규모 조직의 아웃바운드 마케팅 메시지 중 30%가 합성으로 생성되리라고 전망합니다.[1]

생성형 AI 기술에 대한 인식은 전례 없는 속도로 확산되고 있습니다. 미국 내 대규모 조직에서 약 4,500명의 직원을 대상으로 설문 조사를 실시한 Fishbowl에 따르면, 25%가 이미 업무 관련 작업을 보조하는 생성형 AI를 사용하고 있다고 답했습니다.[2]

[1] 'Beyond ChatGPT: The Future of Generative AI for Enterprises', Gartner, 2023년 1월
[2] 'Rate of generative AI adoption in the workplace in the United States 2023, by Industry', Statista, 2023년 12월

기회 포착

현재 다양한 산업의 조직이 생성형 AI가 제시하는 경제적 기회를 잡고자 경쟁을 벌이고 있습니다. 선도적인 금융 기관의 예측이 실현되면, 생성형 AI의 부상은 세계 경제의 새로운 시대를 열 것입니다.

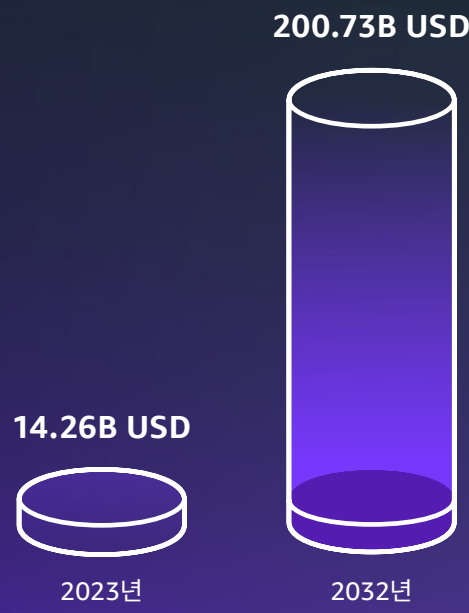
Goldman Sachs의 연구에 따르면, 생성형 AI는 향후 10년 동안 세계 GDP를 최대 7%(약 7조 USD) 끌어올릴 것으로 전망됩니다.[3]

소비자의 관심만으로 이렇게 긍정적인 경제 전망이 도출된 것은 아닙니다. 비즈니스 생산성과 성과를 개선하는 생성형 AI의 잠재력을 생각해보면 AI 기술을 둘러싼 관심과 열정이 이해가 됩니다.

생성형 AI는 모든 규모의 비즈니스와 여러 산업에서 상당한 가치를 창출하기 시작했고, 비즈니스 지형을 근본적으로 뒤흔들 힘을 지닌 혁신적 기술입니다.

글로벌 생성형 AI 시장

CAGR 34.2% 성장 예측[4]



[3] "Generative AI could raise global GDP by 7%", Goldman Sachs, 2023년 4월
[4] 'Generative AI Market Share, Size, Trends, Industry Analysis Report, By Component(Software and Services); By Technology; By End-Use; By Region; Segment Forecast, 2024 - 2032'. Polaris Market Research, 2024년 1월



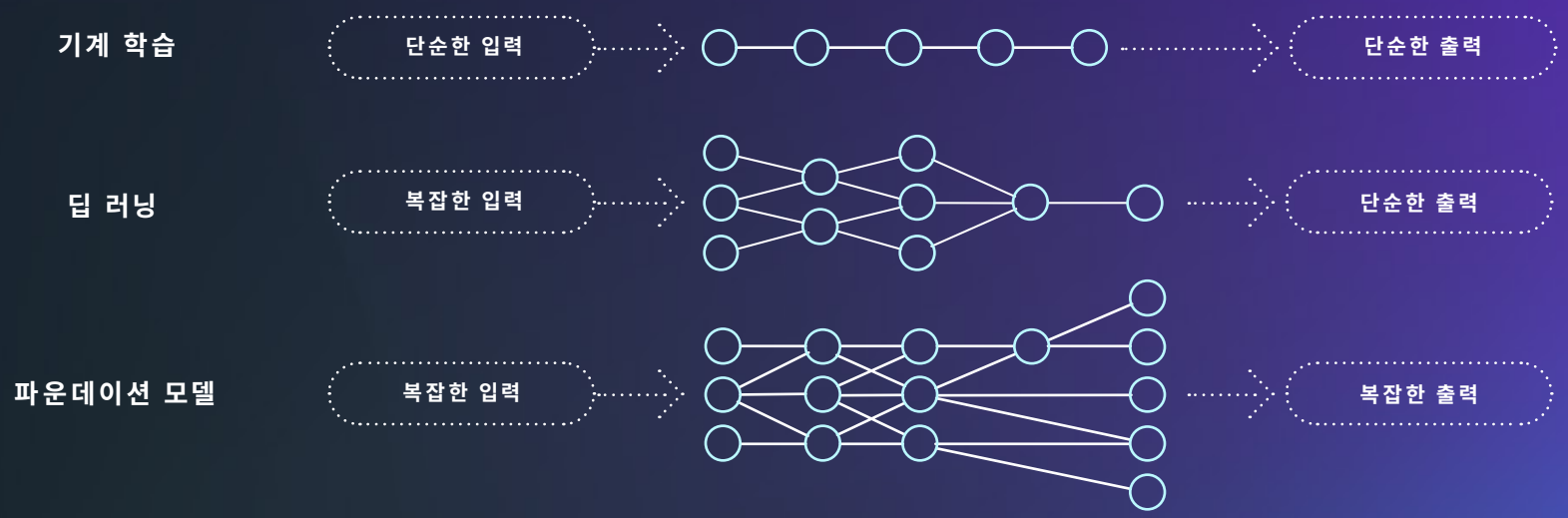
전략적 중요성

전 세계의 모든 산업 조직이 생성형 AI를 활용하여 고객 경험을 재구상하고, 직원 생산성을 높이고, 창의성을 발휘하고, 비즈니스 프로세스를 최적화하고자 합니다.

그러나 대부분의 기업에서는 이러한 이점을 얻을 수 있는 길을 확실히 알지 못합니다.

많은 조직이 경쟁사에서 우위를 점하기 전에 신속하게 생성형 AI에 투자해야 할 필요성에 공감합니다. 그러나 기술 도입, 용도, 결과 실현 및 측정 전략을 제대로 개발한 조직은 거의 없습니다.

이 개요 문서를 읽고 생성형 AI의 비즈니스 가치를 실현할 수 있는 방법을 살펴보세요. 이를 바탕으로 시장의 발전 속도에 발맞추고 경쟁에서 앞서나갈 수 있습니다.



생성형 AI의 이해

생성형 AI의 비즈니스 가치를 완전히 실현하려면 먼저 AI 기술의 작동 원리를 근본적으로 이해해야 합니다.

생성형 AI는 대화, 스토리, 이미지, 동영상, 음악을 포함하여 새로운 콘텐츠와 아이디어를 생성할 수 있는 알고리즘을 의미합니다. 생성형 AI는 방대한 양의 데이터에 대해 사전 훈련을 마친 초대형 ML 모델을 기반으로 구동되는데, 이 ML 모델을 보통 **파운데이션 모델(FM)**이라고 합니다.

기존 ML에서는 숫자 값과 같은 간단한 입력을 예측 값과 같은 간단한 출력에 매핑했습니다. 딥 러닝의 출현으로, 동영상이나 이미지와 같은 복잡한 입력을 수행하고 비교적 간단한 출력(예: 이미지의 고양이 존재 여부)에 매핑할 수 있게 되었습니다. 생성형 AI를 사용하면 대량의 복잡한 데이터를 활용하여 보다 진보한 방식으로 정보를 포착하고 제공할 수 있습니다. 즉, 긴 문서를 요약하여 핵심 정보를 추출하는 등 복잡한 입력을 복잡한 출력에 매핑할 수 있게 된 것입니다.

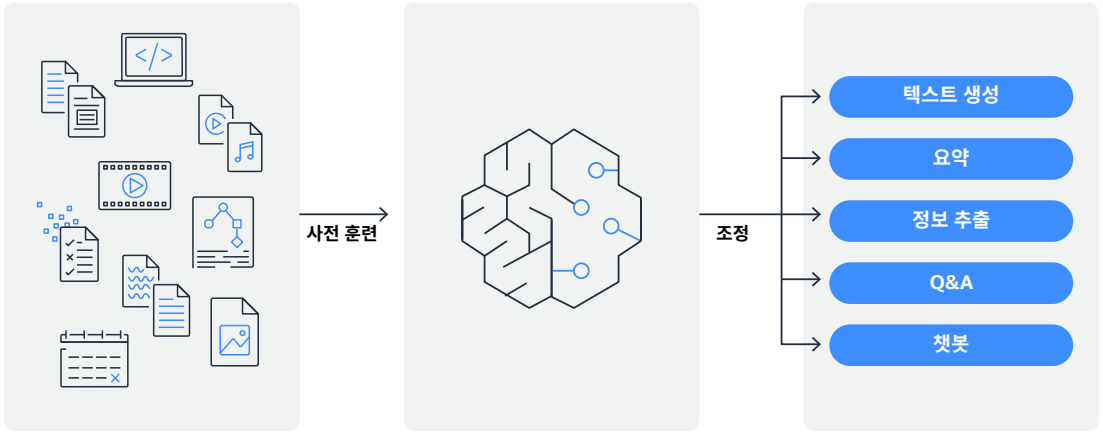
텍스트 기반 생성형 AI 시스템은 **대규모 언어 모델(LLM)**이라는 특정 유형의 FM을 사용합니다. LLM은 다양한 분야를 넘나들며 광범위한 작업을 수행할 수 있습니다. 예를 들어, 코드를 작성하거나, 수학 문제를 풀거나, 대화에 참여하거나, 문서의 정보를 분석하여 질문에 답할 수 있습니다.

데이터를 차별화 요소로 만들기

조직의 비즈니스 요구 사항에 맞는 생성형 AI 애플리케이션을 구축하고자 한다면 조직의 데이터는 전략적 자산입니다. 조직 내 독점 데이터로 FM을 사용자 지정하고 미세 조정하여 '즉시 사용 가능한 FM'에 비해 보다 차별화된 경험을 제공할 수 있습니다. 예를 들어, 쇼핑객 선호도를 추적하는 대형 식료품 체인이라면 FM을 사용자 지정하여 경쟁사의 제품과 완전히 차별화된 더 나은 추천 엔진을 제작할 수 있습니다.

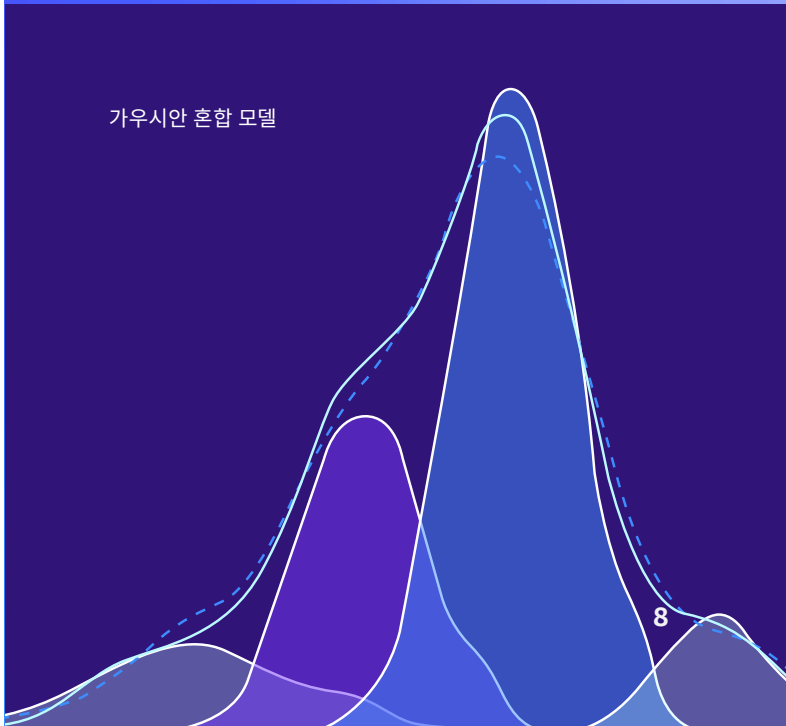
또한, 조직은 사용자 지정 FM을 활용하여 브랜드 특유의 분위기와 스타일을 구현하는 고유 콘텐츠를 쉽게 만들 수도 있습니다. 가령, 내부 배포용 일일 활동 보고서를 자동 생성해야 하는 금융 회사는 과거 보고서를 포함하는 독점 데이터로 FM을 사용자 지정할 수 있습니다. 그런 다음 FM은 이러한 보고서를 읽는 방법과 보고서를 생성하는 데 사용한 데이터를 학습하여 조직의 요구 사항이 더욱 잘 반영된 보고서를 제공합니다.

이제 기술의 작동 원리를 대략적으로 알아보았으니, 조직에서 생성형 AI를 어떻게 업무에 활용할 수 있을지 살펴보겠습니다.



생성형 AI의 역사적 순간:

생성형 AI 애플리케이션을 만드는 데 사용되는 오늘날의 FM은 AI 혁신의 오랜 역사를 토대로 구축되었습니다. 생성형 AI 기능을 갖춘 초기 모델 중 2가지는 1950년대에 개발된 은닉 마르코프 모델(HMM)과 가우시안 혼합 모델(GMM)입니다. HMM은 알려진 데이터를 사용하여 알려지지 않은 데이터를 합리적으로 추측합니다(예: 결과를 기반으로 카드 플레이어가 부정행위를 하는지 예측). GMM은 데이터 그룹(예: 음악 재생 목록)과 해당 데이터 내의 하위 그룹(예: 장르)을 검사하여 알려지지 않은 정보(예: '이 노래는 랩입니다')를 추론합니다. 이 두 모델은 지금도 사용됩니다.



생성형 AI의 비즈니스 기능

다양한 산업의 조직이 생성형 AI를 사용하여 여러 방법으로 생산성을 높이고 비즈니스 가치를 창출합니다.

코드 생성
AI 코딩 동반자 **Amazon Q Developer**로 개발자 생산성 57% 향상[5]

가상 비서
인간과 유사한 대응으로 고객 경험 향상

고객 센터 분석
고객 통화에서 인사이트 요약 및 추출

대화형 검색
회사의 모든 정보에서 인사이트 추출

개인화
개인화된 추천 개선 및 맞춤형 콘텐츠 생성

콘텐츠 생성
텍스트, 이미지, 동영상 및 음악 생성

디자인과 창의성
제안 생성, 프로토타입 생성 및 혁신적인 개념 탐색

문서 처리
문서에서 데이터 및 인사이트 자동으로 추출 및 요약

[5] Amazon이 Amazon CodeWhisperer 평가판 중 수행한 '생산성 챌린지'에서 수집한 데이터

생성형 AI의 역사적 순간:

또 다른 초기 생성형 AI 모델로는 MIT 교수가 1964~1966년에 개발한 챗봇(구 명칭: '채터봇')인 ELIZA가 있습니다. 영화 Pygmalion과 My Fair Lady의 주인공 'Eliza Doolittle'과 이름이 같은 ELIZA 프로그램은 인간의 상호 작용으로부터 '배우며' 더욱 정교해졌습니다. ELIZA의 가장 유명한 사용 사례는 초기 정신 의학 인터뷰를 수행하는 치료사의 행동을 모방하는 것입니다(사용자는 환자 역할).



생성형 AI를 위한 비즈니스 고려 사항

조직에 가장 유용한 생성형 AI의 기능을 식별하고 비즈니스 프로세스에 구현하는 전략을 수립하면서, 생성형 AI 사용 사례에 가장 적합한 FM을 결정해야 합니다.

FM을 지원하기 위해 사용할 인프라도 신중히 고려해야 합니다. 성능 요건을 충족하는 경제적인 인프라는 비즈니스에서 사용하는 모델에 도움이 됩니다.

생성형 AI 애플리케이션 생성에 사용되는 FM을 평가할 때는 다음을 제공하는 모델을 찾아보세요.

1. 보안 및 개인정보 보호 기능이 내장된 생성형 AI 애플리케이션을 쉽게 구축하고 확장할 수 있는 방법
2. 자체 모델을 훈련하고 대규모로 추론을 실행할 수 있는 고성능의 저비용 인프라
3. 업무 수행 방식을 혁신하는 생성형 AI 기반의 애플리케이션
4. 차별화 요소로서의 데이터



책임 있는 AI, 보안 및 개인정보 보호

방대한 크기와 개방형 특성을 가진 FM은 개발 주기 전체에서 정확성, 공정성, 지식 재산(IP) 고려 사항, 할루시네이션, 유해성, 개인정보 보호와 같은 책임 있는 AI 관련 우려를 정의, 측정 및 완화하는 데 있어 새로운 문제를 제기합니다. 예를 들어, 공정성의 문제를 살펴보겠습니다. 우리는 의사를 지칭할 때 남성 대명사와 여성 대명사를 같은 비율로 사용하라고 LLM에 요청할 수 있을까요? 프롬프트에서 의사가 수염이 있다고 설명하는 경우에도 동일하게 적용해야 할까요? 그리고 다른 직업의 경우도 마찬가지일까요? 미국여자프로농구(WNBA)는 어떨까요? LLM의 맥락에서 공정성을 단순히 정의하기란 어렵습니다. 그렇기에 새로운 접근 방식과 솔루션이 필요합니다.

생성형 AI 기술과 그 기술의 사용법이 계속 진화하면서, 새로운 문제가 발생하여 관심과 완화를 위한 노력이 추가로 필요할 것입니다. 이러한 과제를 해결하고 혁신을 촉진하고자 **학계, 산업계, 정부 파트너**가 함께 협력하여 생성형 AI가 책임감 있고 개인정보를 보호하는 안전한 방식으로 계속 진화할 수 있도록 새로운 솔루션과 개념을 모색하고 있습니다.

데이터 개인정보 보호와 보안은 생성형 AI를 책임감 있게 확장하는 데 매우 중요합니다. 모델을 사용자 지정하고 미세 조정할 때 조직은 데이터의 사용처와 사용 방법을 파악할 필요가 있습니다. 보호 대상 데이터가 공개 모델 훈련에 사용되지 않으며, 고객 데이터가 비공개로 유지될 것이라는 믿음이 요구됩니다. 조직은 비즈니스 애플리케이션에 적합한 보안, 확장성 및 개인정보 보호 기능을 처음부터 기본으로 제공해야 합니다.

블로그 게시물 **Responsible AI in the Generative Era** 보기
더 알아보기 ›

생성형 AI의 역사적 순간:

2014년, 최초의 생성적 적대 신경망(GAN)이 개발되면서 생성형 AI는 큰 돌파구를 찾았습니다. GAN은 두 모델('생성자'와 '구분자')이 경쟁하는 제로섬 게임입니다. 생성자는 점점 더 '진짜'로 보이는 콘텐츠를 제조하는 반면, 구분자는 상대의 기술을 분석하여 가짜를 더 효과적으로 가려냅니다. AI의 훈련에 다른 AI를 사용하는 참신한 접근법은 혁신 그 자체였으며, GAN은 이렇게 디지털 이미지의 새로운 시대를 열었습니다.

다음 섹션에서는 클라우드 모범 사례, 문화 변화, 조직 민첩성, 생성형 AI 혁신에 대한 경험과 전략을 공유하는 AWS 업계 리더를 소개합니다.

생성형 AI에 대한 경영진 인사이트

주류 기술로 전환하는 모든 신기술이 그렇듯이, 생성형 AI는 시간이 지나면서 학습 효율이 높아집니다. 경영진은 생성형 AI가 적합한 분야와 그렇지 않은 분야, 효과적인 활용 방법, 비기술적 고려 사항을 알고 싶어 합니다.

역사를 되짚어 보면 기술은 사람, 과정, 기량, 문화와 함께할 때만 심오하고 긍정적인 변화를 가져왔습니다. 이러한 교훈과 더불어 수천에 달하는 AWS 고객들에게 얻은 정보를 바탕으로, 생성형 AI에 관심이 있는 분들에게 다음과 같이 간단한 조언을 드리겠습니다.

먼저 호기심을 가져야 합니다. 생성형 AI가 무엇인지, 사람들의 상상력을 어떻게 자극했는지, 어떤 문제를 해결할 수 있는지 알아보세요. 모델을 사용자 지정할 때는 데이터 보안과 같은 영역을 자세히 살펴보고 관련 작업을 IT 팀에만 맡기지 말고 직접 알아보도록 장려해야 합니다.

둘째로, 생각의 폭을 넓히고 고객의 입장을 고려하는 것부터 시작하세요. 이것이 AWS의 사고방식입니다. 공급망 효율성을 개선하거나, 새로운 서비스를 개발하거나, 고객 서비스를 강화할 수 있는 기회를 포함한 여러 비즈니스 기회를 제대로 이해해야 합니다. 솔루션으로 넘어가기 전에 비용을 절감하거나, 복원력을 높이거나, 전체 매출을 증대할 수 있는 부문을 찾아 문제를 먼저 알아보세요. 그리고 기회를 야심차게 추구해야 합니다. 소박한 기회를 추구하면 소박한 미래가 실현됩니다.

마지막으로 지금 시작하세요. 대부분의 비즈니스 이니셔티브는 동력을 얻기까지 시간이 걸리므로, 신속하게 실험을 시작해야 합니다. 기업 대부분이 기술을 채택할 때 으레 그러하듯, 끊임없이 계획하고 완벽한 시기를 기다리기보다 부딪히면서 더 많이 배울 수 있습니다.



Phil Le-Brun
AWS Enterprise Strategy Director

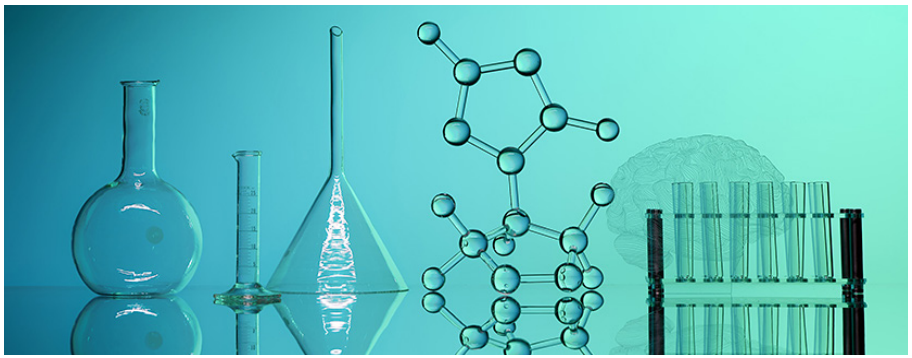
2019년 AWS에 입사한 Phil은 McDonald's Corporation에서 25년 동안 근무하며 디지털 트랜스포메이션을 주도한 경험을 비롯해 대규모로 기술을 구현하면서 실질적인 교훈을 얻었습니다. Phil은 기업 고객에게 지식을 공유하여 조직 민첩성 지원, 고객 중심주의 강화 등 클라우드 기반 기술 목표를 달성할 수 있도록 지원합니다.

생성형 AI의 일반적인 산업별 사용 사례

이전 섹션에서는 모든 비즈니스에서 활용할 수 있는 생성형 AI의 기능과 비즈니스 요구에 적합한 모델, 인프라를 선택하는 방법을 알아보았습니다. 이제, 좀 더 깊이 파고들어 구체적인 산업별 생성형 AI 사용 사례를 살펴보도록 하겠습니다.

의료 서비스 및 생명 과학

- **제약 연구 및 개발(R&D) 혁신 가속화:** 생명과학 조직은 더욱 효과적인 신약 후보의 개발을 더 효율적으로 가속화할 수 있습니다. 예를 들어, 연구원은 단백질 구조를 예측함으로써 생물학적 표적을 더 잘 식별하고, 새로운 아미노산 서열을 생성하고, 특정 표적의 도킹 위치를 식별하여 정밀 치료제를 만들 수 있습니다. 이렇게 되면 궁극적으로 더 많은 표적화를 수행하고, 환자에게 효과적인 치료 옵션을 제공하게 됩니다.



- **임상 참여 및 환자 경험 개선:** 생성형 AI를 사용하여 환자 탐색, 치료 조정, 의료 서비스 제공자의 생산성을 개선할 수 있습니다. 대화형 상호 작용을 통해 임상의는 전자 건강 기록, 연구 간행물, 의료 정책 데이터에 보다 쉽게 액세스할 수 있게 됩니다. 생성형 AI는 또한 지능형 문서 처리(IDP)를 용이하게 하고, 맞춤형 의료/약리학 업무 지침을 제공하고, 신약 개발의 안정성과 효능을 보장하기 위한 효과적인 임상 시험 프로토콜을 설계하는 데 도움을 주어 디지털 트랜스포메이션과 개인화된 치료를 가속화할 수 있습니다.

- **건강 및 과학 데이터 요약:** 연구원과 임상의는 생성형 AI를 활용하여 인사이트 확보 시간을 줄이고, 데이터 중심 의사 결정에 필요한 데이터를 신속하게 찾을 수 있습니다. 의료 서비스의 경우 차트 노트 구성 요소 및 노트 요약을 자동 생성하고, 워크플로를 간소화하며, 관리 작업을 자동화함으로써 검색 및 요약을 통해 효율성을 높일 수 있습니다. 임상 시험의 패턴 식별, 연구 논문 수십만 개의 조사 결과를 요약할 수 있는 기능, 가치 사슬 전반의 효율성 파악 등 생명과학 부문의 응용 분야는 무궁무진합니다.



금융 서비스

- **경험 개선:** 금융 서비스 기업은 요약과 쿼리를 통해 고객 상호 작용 리포지토리를 활용하는 등 문제를 빨리 해결하고 제품과 권장 사항을 개인화하는 챗봇을 배포하여 고객과 직원에게 제공하는 서비스를 개선할 수 있습니다. AI는 또한 여러 내부 소스를 바탕으로 기술 자료 문서를 만들 수 있습니다.
- **지식 근로자 효율성 향상:** 금융 서비스 기업의 지식 근로자는 투자 조사, 대출 문서, 보험 정책, 규제 커뮤니케이션, 정보 요청서(RFI), 비즈니스 서신 초안을 신속하게 작성하고, 깊이 있는 고객 행동 인사이트를 얻고, 협업을 개선하고, 비정형 콘텐츠의 비즈니스 가치를 높일 수 있습니다.
- **시장 심리 분석:** 금융 서비스 기업은 소셜 미디어, 뉴스 기사 및 금융 데이터를 신속하고 철저하게 분석하여 시장 의견을 표면화하고 기회를 빨리 파악하고 위험을 사전에 완화할 수 있습니다.
- **신제품 개발 및 비즈니스 프로세스 자동화:** 금융 서비스 기업은 대규모 비정형 데이터 소스에서 온디맨드 정형 데이터 제품을 생성할 수 있습니다. 생성형 AI는 최종 사용자가 활용하는 생산성 도구를 개선하고 자동 코드 생성을 통해 개발자 생산성을 높여줍니다.



자동차 및 제조

- **제품 디자인 향상:** 제조업체는 AI를 이용해서 기계 부품의 디자인을 최적화하거나, 새로운 재료, 칩, 부품 디자인을 만들 수 있습니다. 이를 통해 품질과 내구성이 높아지고, 가격이 낮아지고, 생산 공정이 간소화됩니다.
- **새로운 차량 내 경험 조성:** 가상 비서와 맞춤형 경로 추천은 운전자와 승객의 경험을 개선할 수 있습니다.
- **테스트와 유지 보수 개선:** AI는 데이터 시트에서 누락된 정보를 생성하여 제품 테스트를 개선하고, 새로운 지원 유지 보수 사례를 활용하여 소비자가 사용 중인 제품을 포함한 기계를 효율적으로 유지 보수 및 서비스할 수 있습니다.
- **공장의 전체적인 장비 효율성 향상:** 과거 기계 유지 보수 데이터, 수리 데이터, 장비 매뉴얼, 생산 데이터 및 기타 제조업체의 잠재적인 데이터를 디지털화하고 캡처하여 유지 보수, 수리 또는 장비 파라미터에 대한 제안을 생성하고 생산성, 가용성, 품질을 개선합니다.



교육

- **텍스트 요약:** 학생과 교사는 연구 문서, 강의 기록 및 강의 노트를 간결하게 요약하여 검색하고 탐색하기 쉽게 만들 수 있습니다.
- **자동화 개선:** AI는 정보를 샘플 시험 문제로 변환하고, 채점을 가속화하고, 광범위한 요소에 걸쳐 학생의 성과를 측정하며, 교사 및 학생에게 맞춤형 피드백 및 권장 사항을 제공할 수 있습니다.
- **맞춤형 학습 환경:** 교육자는 학생 단체 또는 개별 학생을 위한 맞춤형 학습 계획을 만들고 시뮬레이션과 가상 현실을 활용하여 학습 몰입도를 높일 수 있습니다.



미디어 및 엔터테인먼트

- **콘텐츠 제작 촉진:** 콘셉트 디자인과 스토리보드부터 포스트 프로덕션 워크플로까지, 미디어 및 엔터테인먼트(M&E) 기업은 하위 작업을 자동화하여 프로덕션 속도를 높이고 크리에이티브 담당자의 반복 작업 속도를 높여 디렉터의 비전을 실현할 수 있습니다.
- **음악 발전:** 아티스트는 AI 생성 음악으로 앨범을 보완하고 발전시켜 새로운 장르를 만들 수 있습니다.
- **미디어 공급망 개선:** 생성형 AI 애플리케이션은 현지화, 콘텐츠 조정, 홍보 자료의 아트워크 제작, 콘텐츠 복원과 같은 작업을 지원하거나 자동화할 수 있습니다.

생성형 AI의 역사적 순간:

2017년, 새로운 딥 러닝 모델 유형인 트랜스포머의 도입은 현대적 생성형 AI의 발판을 마련했습니다. 입력 데이터를 분해하여 처리한 후 재조립하는 기존 모델과 달리 트랜스포머는 전체 입력을 한 번에 처리합니다. 따라서 입력 내용의 전체 맥락을 이해하는 것이 무엇보다 중요한 자연어 처리(NLP)에 이상적입니다. 2018년, OpenAI는 최초의 사전 훈련된 생성형 트랜스포머(GPT)를 만들어 기술을 한층 발전시켰습니다. 이를 토대로 OpenAI는 2019년에 GPT-2 엔진을 개발하여 2022년 말에 출시된 ChatGPT 제작에 사용했습니다.

AWS로 생성형 AI의 성공을 이끄는 비결

AWS와 함께하면 조직에서 생성형 AI의 온전한 비즈니스 가치를 실현할 수 있습니다. 애플리케이션을 재구성하고, 새로운 형태의 고객 경험을 만들고, 전례 없는 수준으로 생산성을 높이고, 궁극적으로 비즈니스를 혁신하세요.

경험과 전문성

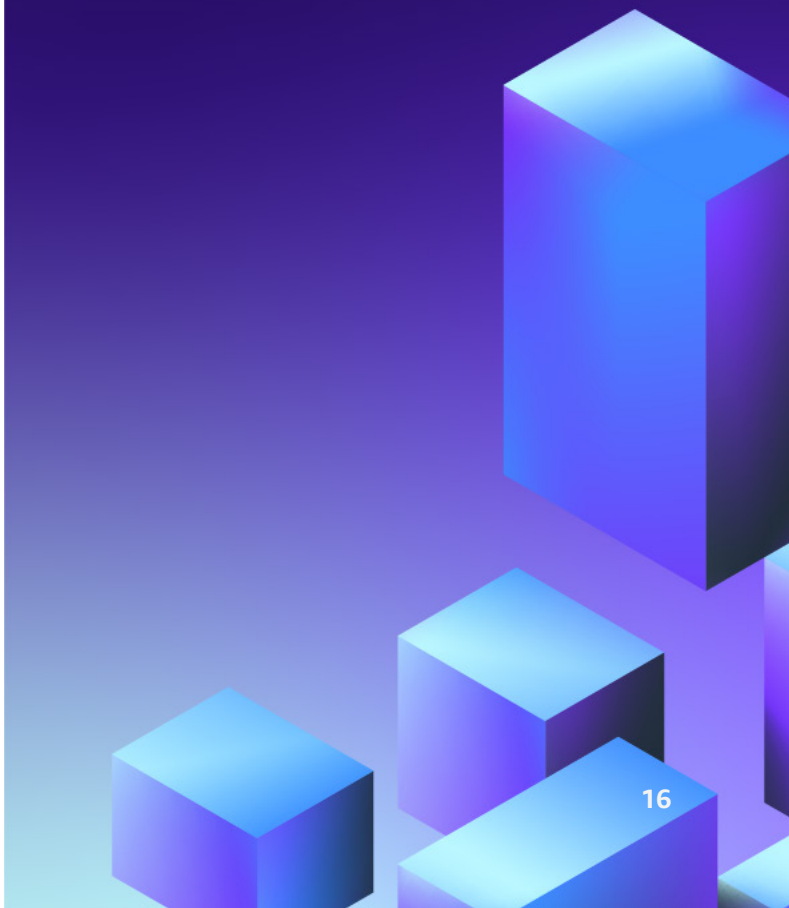
AWS의 핵심 장점 중 하나는 20년 이상의 집중 투자를 통해 쌓아온 풍부한 AI 자산입니다. 실제로 10만이 넘는 고객이 현재 AI 기술과 ML 도구를 위해 AWS를 사용하고 있습니다.

AWS를 뒷받침하는 Amazon은 ML 기능을 활용하여 전자 상거래 추천 엔진을 구동하고, 주문 처리 센터에서 로봇 픽업 경로를 최적화합니다. 나아가 ML은 Amazon의 공급망, 예측 및 용량 계획을 알려줍니다.

딥 러닝은 Amazon Prime Air 드론 배송 시스템과 Amazon Go 컴퓨터 비전(CV) 기술에도 적용되어 고객이 기존의 결제 과정 없이 물건을 골라 매장을 나가는 혁신적인 소매 경험을 선사합니다. 또한, 30개가 넘는 여러 ML 시스템에서 사용할 수 있는 Alexa는 매주 수십억 차례의 작업을 수행하며 고객을 지원합니다.

수천 명의 전담 ML 엔지니어를 보유한 Amazon과 AWS에서 생성형 AI와 ML은 기업 문화에 깊숙이 통합되어 미래를 만들어 나가고 있습니다.

현재 10만 이
넘는 고객이 AI 기술과 ML 도구를 위해
AWS를 사용하고 있습니다.



AWS로 구축해야 하는 이유

다양한 형태와 규모의 조직이 여러 이유로 AWS에서 생성형 AI 및 ML 애플리케이션을 구축합니다. 고객의 말을 빌리자면, AWS를 기반으로 구축할 때 이점은 다음과 같습니다.

보안 및 개인정보 보호 기능이 내장된 생성형 AI 애플리케이션을 구축하고 확장할 수 있는 가장 쉬운 방법

Amazon Bedrock은 단일 API를 통해 AI21 Labs, Anthropic, Cohere, Meta, Mistral AI, Stability AI 및 Amazon과 같은 주요 AI 회사의 고성능 FM을 제공하는 완전관리형 서비스입니다. 포괄적인 생성형 AI 프레임워크인 Amazon Bedrock은 보안, 개인 정보 보호 및 책임 있는 AI를 사용하여 생성형 AI 애플리케이션을 구축하는 데 필요한 광범위한 도구도 제공합니다. 사용 사례에 가장 적합한 주요 FM을 손쉽게 실험 및 평가하고, 미세 조정 및 검색 증강 생성(RAG)과 같은 기술로 자체 데이터를 활용해 비공개로 사용자 지정하고, 엔터프라이즈 시스템 및 데이터 소스를 이용하여 작업을 실행하는 에이전트를 구축할 수 있습니다. Amazon Bedrock은 서버리스이므로, 인프라를 관리할 필요가 없으며, 이미 익숙한 AWS 서비스를 사용하여 생성형 AI 기능을 애플리케이션에 안전하게 통합 및 배포할 수 있습니다. Amazon Bedrock의 차별화된 기능을 통해 조직은 성장 목표에 맞춰 안전하고 미래 지향적인 AI 전략을 유지하면서 새로운 수익 기회를 창출하고, 운영 효율성을 높이며, 탁월한 고객 경험을 제공할 수 있습니다.

생성형 AI를 위한 고성능의 저비용 인프라

AWS는 수년 동안 생성형 AI 워크로드에 대해 최고 수준의 성능과 비용 최적화를 제공하는 실리콘 개발에 투자해 왔습니다. 그 결과로 탄생한 **AWS Trainium**과 **AWS Inferentia**는 클라우드에서 모델 훈련과 추론 실행을 가장 저렴한 비용으로 제공합니다. AWS가 개발한 **Amazon Elastic Compute Cloud**(Amazon EC2) 인스턴스로 인해 사용자는 모델 훈련과 추론 실행 기능을 이용할 수 있습니다. 예를 들어, Trainium을 기반으로 한 **Amazon EC2 Trn1** 인스턴스는 훈련 비용을 최대 50% 절감하며[6], AWS Inferentia2를 기반으로 한 **Amazon EC2 Inf2** 인스턴스는 추론당 비용을 최대 40% 줄여줍니다[7].



[6] 동급 Amazon EC2 인스턴스 대비
[7] 이전 세대 AWS Inferentia 기반 인스턴스 대비

차별화 요소로서의 데이터

AWS를 사용하면 쉽게 조직의 데이터를 전략적 자산으로 사용하여 FM을 사용자 지정하고 더욱 차별화된 경험을 구축할 수 있습니다. 일반 생성형 AI 애플리케이션과 비즈니스와 고객을 진정으로 이해하는 애플리케이션을 가르는 차이점은 데이터에서 나옵니다. 또한 가장 포괄적인 데이터 및 AI 서비스 세트로 자체 데이터를 사용하여 AWS에서 FM을 안전하게 사용자 지정하고 비즈니스, 데이터 및 고객에 전문화된 모델을 구축할 수 있습니다.

업무 수행 방식을 혁신하는 기업용 생성형 AI 기반의 애플리케이션

AWS는 고객이 생성형 AI를 사용하여 작업을 완료하는 방식을 혁신하는 강력한 신규 애플리케이션을 구축하고 있습니다. **Amazon Q Developer**에서 기업의 코딩을 간소화하고, **Amazon Q in QuickSight**에서 비즈니스 인텔리전스(BI)를 간소화하고, **AWS HealthScribe**에서 의료 기관의 임상 효율성을 개선하는 목적별 대화형 에이전트로 생산성을 강화합니다. 보안, 개인 정보 보호, 책임 있는 AI를 가장 우선시하고 기존 데이터 소스 및 애플리케이션으로 쉽게 사용자 지정 및 통합할 수 있어 기업은 번거로운 작업 없이 생성형 AI를 신속하게 활용할 수 있습니다.

책임 있는 AI 관련 추가 리소스:

AWS의 책임 있는 AI 리소스 허브 ›

eBook: 대중화, 운영화, 책임성: 성공적 AI 성과의 3대 요소 ›

AWS 생성형 AI 서비스

다음을 포함한 여러 AWS 기술을 통해 생성형 AI 애플리케이션을 지원합니다.



Bedrock으로 지속 가능한 성장 실현: 생성형 AI 애플리케이션을 구축하고 확장하는 가장 쉬운 방법

Bedrock은 다양한 업계의 여러 사용 사례를 지원하기 위해 업계 최고의 다양한 FM을 제공합니다. 모든 사용 사례에 적합한 한 가지 모델은 없습니다. 모델은 기능, 가격 및 성능에 따라 다릅니다.

Amazon Bedrock 모델 선택에 대해 자세히 알아보기 ›

AWS Trainium: 비용을 최대 50% 절감하면서[8] 모델 훈련 시간을 단축하는 데 이 ML 모델 액셀러레이터를 사용합니다.

AWS Inferentia2: 추론당 비용을 최대 40% 절감하면서[9] 고성능 FM 추론을 실행하는 액셀러레이터입니다.

Amazon Q Developer: 개인 개발자에게 무료로 제공되는 AI 코딩 도우미로, 보안을 강화하면서 애플리케이션 개발 속도를 57% 높여줍니다[10].

Amazon Q Business: 엔터프라이즈 시스템의 데이터 및 정보를 기반으로 질문에 답하고, 요약을 제공하고, 콘텐츠를 생성하며, 작업을 완료할 수 있는 생성형 AI 기반 어시스턴트입니다.

Amazon QuickSight 생성형 BI: Amazon QuickSight의 생성형 BI 기능을 사용하여 기존의 다단계 BI 작업을 직관적이고 강력한 자연어 경험으로 전환합니다.

Amazon SageMaker: 관리형 인프라와 도구로 자체 FM을 구축하여 확장 가능하고, 신뢰할 수 있으며, 안전한 모델의 구축, 훈련, 배포를 가속화하세요.

Amazon SageMaker JumpStart: ML을 빠르게 시작할 수 있도록 알고리즘, 모델 및 ML 솔루션에 대한 액세스를 제공하는 ML 허브입니다. SageMaker JumpStart를 사용하면 ML 실무자가 **공개적으로 사용 가능한 다양한 FM** 중에서 선택할 수 있습니다. ML 실무자는 네트워크 격리 환경의 전용 SageMaker 인스턴스에 FM을 배포하고, 모델 훈련과 배포를 위해 SageMaker를 사용하여 모델을 사용자 지정할 수 있습니다.

[8] AWS Trainium은 동급 Amazon EC2 인스턴스에 비해 훈련 비용을 최대 50% 절감
[9] AWS Inferentia는 동급 Amazon EC2 인스턴스에 비해 추론당 비용을 최대 40% 절감
[10] Amazon이 Amazon CodeWhisperer 평가판 중 수행한 '생산성 챌린지'에서 수집한 데이터

다음 단계

이제 생성형 AI, 생성형 AI가 할 수 있는 일, 잠재적인 비즈니스 이점을 깊이 있게 이해했으니, 다음 단계로 목표를 명확히 정의하고 생성형 AI를 활용하는 사용 사례를 파악해야 합니다. 소규모로 실험하고, 단순하되 정확한 목표를 정해 시작하는 것이 가장 좋습니다. 빠르게 결과를 달성했다면 상위 직무와 외부로 활동 규모를 확장하면 됩니다.

전문가와 협업하여 데이터 가용성, 데이터 품질, 생성형 AI와 관련된 윤리적 영향과 같은 요소를 고려하기를 적극 권장합니다. 나아가 비용, 확장성, 에너지 소비에 상당한 영향을 미칠 수 있는 인프라도 사전에 고려해야 합니다. AWS 전문가와 협력하면 의사 결정 프로세스와 구현 단계 전반에 걸쳐 귀중한 지침을 얻을 수 있습니다.

지금이 중요한 시점

생성형 AI의 극적인 성장으로 우리는 중대한 분기점에 서게 되었습니다. FM은 매일 더 정교해지고 강력해지고 있습니다. 조직에는 새로운 고객 경험을 창출하고, 전례 없는 수준의 효율성과 생산성을 이끌어 비즈니스를 혁신할 수 있는 힘이 됩니다.

따라서 엄청난 기술 발전이 이루어지는 새 시대에서 경쟁하려면 모든 조직이 생성형 AI를 혁신 로드맵에 포함하는 것을 고려해야 한다는 확고한 결론에 도달하게 됩니다.

생성형 AI를 위한 가장 비용 효율적인 클라우드 인프라, 다수의 AI 제품/서비스/솔루션, 오랫동안 신뢰받는 AI 전문성을 갖춘 AWS는 생성형 AI의 가능성을 성과로 바꾸는 데 도움을 줄 수 있습니다.



지금 AWS와 협력하여 생성형 AI 여정의 속도를 높여보세요.

시작하기 ›

© 2024, Amazon Web Services, Inc. 또는 자회사. All rights reserved.