

엔터프라이즈급 생성형 AI로 확장성을 실현하기 위한 최고 정보 책임자 가이드

최근 연구는 IT 리더들이 AI의 다음 단계를 어떻게 준비하고 있는지에 대한 인사이트를 제공합니다.



생성형 AI는 실험 단계를 넘어 운영 단계로 진입했습니다. 이제 IT 조직은 전사적 실행과 측정 가능한 성과를 지원하기 위해 이 기술을 안전하고 신뢰성 있게 운영해야 합니다. 이러한 변화는 생성형 AI를 예측 가능한 비용, 거버넌스, 비즈니스 연속성, 규제 리스크 및 장기적 경쟁력을 결정짓는 전략적 운영 및 아키텍처 우선순위로 격상시킵니다.

새로운 Foundry Market Pulse 연구에 따르면, 최고 정보 책임자와 기업 리더들은 생성형 AI를 전사적 자원 관리, 고객 관계 관리 및 기타 미션 크리티컬 시스템과 동일한 수준의 엔터프라이즈급 역량으로 간주하기 시작했습니다. 설문에 참여한 거의 모든 고위 IT 리더(96%)는 생성형 AI가 조직의 비즈니스 전략에 중요하다고 인정합니다. 차별화된 경쟁 우위 요소라는 점에 동의하는 비율도 거의 같습니다 (그림 1 참조).

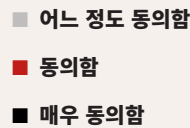
최고 정보 책임자들이 생성형 AI에 대한 방대한 투자가 확장되고, 비즈니스에 적시에 투자 수익을 제공하며, 회복 탄력성과 지속 가능한 성공의 기반을 마련해야 한다는 책임을 지고 있다는 점을 고려할 때, 이러한 전략적 사고방식은 타당합니다. 결과적으로 IT 리더들은 생성형 AI를 하이브리드 환경 전반에서 확장 가능하도록 설계되고, 회복 탄력성, 예측 가능성 및 책임성을 이끄는 동일한 원칙에 따라 관리되는 표준화된 플랫폼으로 운영해야 합니다.

미래 지향적이고 성공적인 리더들은 이 기술을 다른 어떤 비즈니스 핵심 애플리케이션과 동등한 수준으로 간주하며(매우 강력한 것이긴 하지만), 다양한 환경에 걸쳐 배포하고 감독하기 위해 기업 수준의 인프라,

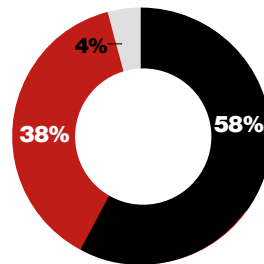
거버넌스 및 전문 역량을 필요로 한다고 봅니다. 이는 기업용 인프라로서 설계되고, 관리되며, 운영되어야 합니다.

그림 1

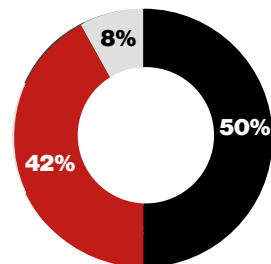
기업의 최우선 과제가 된 생성형 AI



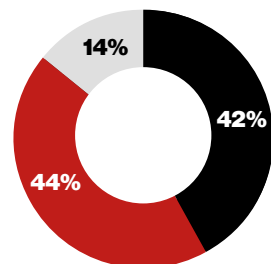
향후 12개월 동안 생성형 AI는 우리 비즈니스 전략에 중요합니다



우리 비즈니스는 생성형 AI를 경쟁 차별화 요소로 봅니다



생성형 AI는 전사적 자원 관리 및 고객 관계 관리와 동등한 수준입니다



출처: FOUNDRY

기업의 최우선 과제가 된 생성형 AI

생성형 AI 도입은 미션 크리티컬 프로덕션 환경 전반에서 빠르게 확산되고 있습니다. 예를 들어, 설문 응답자의 3분의 2가 현재 조직 내 서비스형 소프트웨어(SaaS) 애플리케이션과 기업 업무 프로세스에 생성형 AI를 통합하여 활용 중이라고 답했습니다. 이러한 도입 사례들은 정형 데이터, 비정형 데이터, 객체 데이터를 포괄적으로 활용하며, 이는 사실상 기업 내부를 흐르는 전체 데이터 생태계를 아우르는 것입니다 (그림 2 참조).

생성형 AI 워크플로의 광범위하고 깊이 있게 파고드는 영향력을 보여주는 또 다른 지표는 다음과 같습니다: 대부분의 조직은 퍼블릭 클라우드, 프라이빗 클라우드, 데이터 센터 및 엣지(Edge) 로케이션을 포함한 하이브리드 클라우드 환경 전반에서 생성형 AI 애플리케이션을 운영 중이거나 운영할 계획을 가지고 있습니다. 최고 정보 책임자들이 빠르게 깨닫고 있듯이, 이러한 규모의 확장은 수많은 통합, 보안, 주권 및 성능 관련 과제를 야기합니다.

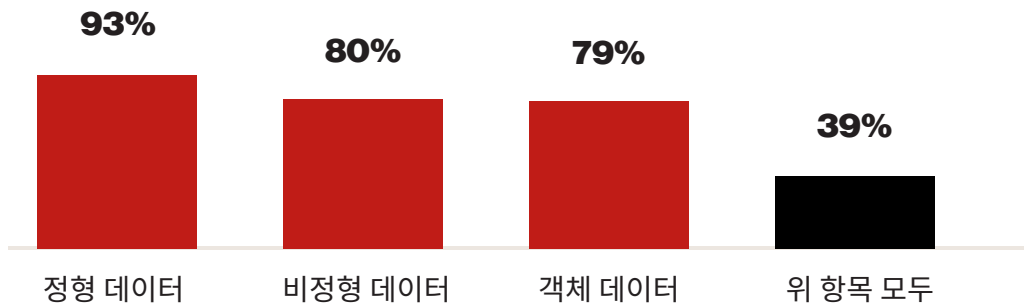
이러한 문제를 극복할 수 있는 IT 리더들에게 그 보상은 충분히 가치가 있을 것입니다.

IT 리더들은 생성형 AI 투자를 통해 다음과 같은 다양한 긍정적인 결과를 기대하고 있습니다:

- 생산성 및 효율성 향상
- 고객 경험 개선
- 매출 증대
- 전사적 비즈니스 프로세스 혁신
- 혁신 및 신제품 개발

하지만 초기 운영 단계에서의 성공이 곧 기업의 준비 상태나 장기적인 규모의 지속 가능성을 의미하는 것은 아닙니다.

Nutanix의 제품 및 솔루션 마케팅 담당 수석 디렉터이자 업계 전문가인 하사 코티 켈라는 다음과 같이 말합니다. "생성형 AI가 기업의 전략적 영역에 어떤 영향을 미칠 수 있는지에 대해 확실히 긍정적인 전망이 있습니다. 그러나 다른 새로운 기술들과 마찬가지로, 많은 조직이 여전히 초기 단계에 머물러 있습니다. 최고 정보 책임자들 사이의 논의는 이제 기대감에서 '우리가 모든 개념 증명(PoC)에서 배운 것을 어떻게 활용하여 확장 가능한 기반 요소를 구축할 것인가?'로 옮겨가고 있습니다."

그림 2 | 생성형 AI는 비즈니스 및 IT 환경 전반에 걸쳐 광범위하게 활용되고 있습니다**생성형 AI 프로젝트에서 접근하는 데이터 유형****프로덕션 환경에서의 생성형 AI 도입.**

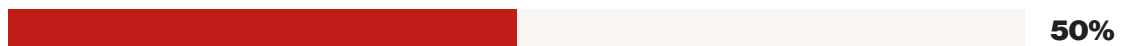
서비스형 소프트웨어(SaaS) 기반 생성형 AI 애플리케이션 운영



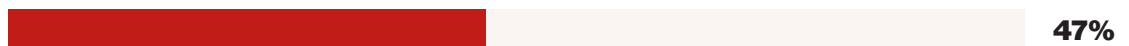
운영 중인 엔터프라이즈 또는 시스템에 내장된 생성형 AI



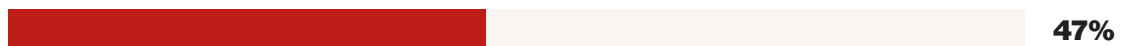
클라우드 호스팅 모델을 통한 퍼블릭 추론



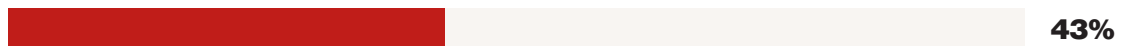
자체 구축 생성형 AI 애플리케이션 운영



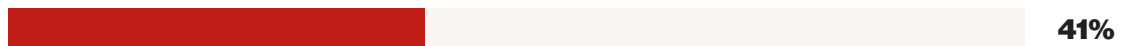
분석 또는 BI 도구와 통합된 프로덕션급 생성형 AI



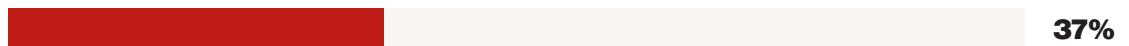
기업 소유/관리 인프라를 사용한 프라이빗 추론



내부 배포를 위해 미세 조정된 파운데이션 모델



실무 환경 내 생성형 AI 사용을 능동적으로 관리하는 거버넌스 및 모니터링 체계



출처: FOUNDRY

과제: 리스크, 인프라 및 규모

AI 확장에 있어 조직이 직면한 인프라 과제에 대한 질문에, 응답자들은 보안 및 규정 준수, 성능 및 안정성, 그리고 데이터 관리 통합을 상위 3가지 과제로 꼽았습니다.

또한, IT 리더들은 생성형 AI 목표를 달성하는 데 있어 다음과 같은 중요한 과제들을 지적했습니다:

■ 데이터 리스크: 대부분의 응답자(94%)

는 생성형 AI 관련 데이터 개인정보 보호 및 규정 준수 리스크를 관리할 수 있는 능력에 대해 최소한 보통 수준 이상의 자신감을 보였습니다. 그러나 이들은 AI의 데이터 거버넌스 요구 사항을 지원하기 위해 플랫폼, 정책 및 도구를 업그레이드하는 데 필요한 노력을 과소평가하고 있을 수 있습니다. 예를 들어, AI에 사용되는 민감한 데이터는 종종 국가 또는 조직의 경계 내에 유지되어야 합니다. 생성형 AI는 데이터, 모델 및 처리가 정의된 관할권 또는 기업의 통제 범위 내에 머물도록 보장하는 주권적 실행을 점점 더 요구하고 있습니다. 그러나 많은 플랫폼은 이러한 요구 사항을 충족할 투명성과 집행 능력이 부족합니다. 또한 조직은 AI 파이프라인을 보호하고 데이터, 모델 및 컴퓨팅 계층 전반에 걸쳐 정책을 시행할 수 있는 통합 거버넌스 체계가 부족할 수 있습니다.

■ 전문 역량 준비도: AI 거버넌스 및 에이

전트 워크플로 설계를 포함한 AI 핵심 기술에 대한 수요가 공급을 앞지르면서, 기업 리더들은 에이전트 기반 시스템의 통합 및 오케스트레이션 요구 사항을 해결할 준비가 되어 있지 않은 상태입니다. 인재 부족은 확장성과 회복 탄력성 측면의 리스크를 초래하기도 합니다. 기술 격차가 벌어짐에 따라, 최고 정보 책임자들은 반복 가능하고 자동화된 AI 인프라와 거버넌스를 구축하여 부족한 전문 AI 인재에 대한 의존도를 줄여야 할 것입니다.

■ 예측 불가능한 비용: IT 리더들은 오랫동안

사용량 기반(pay-as-you-go) 클라우드 서비스의 비용을 예측하고 관리하는 방식을 개선하기 위해 고군분투해 왔습니다. AI는 IT 비용 관리에 훨씬 더 많은 예측 불가능성을 도입합니다. 고도화된 AI 시스템은 토큰(학습 및 추론 과정에서 AI 모델이 처리하는 데이터 단위), 지연 시간 및 에너지 측면에서 막대한 추가 비용을 발생시킬 수 있습니다. 파편화된 솔루션은 복잡성을 증가시키고 가시성을 떨어뜨려 중복과 낭비를 초래할 수 있습니다. 예측 불가능한 비용은 최고 정보 책임자들로 하여금 하이브리드 환경 전반에 걸쳐 AI 도구 및 플랫폼에 대한 가시성, 최적화 및 비용 추적 기능을 개선하도록 이끌 것입니다.

그림 3

IT 리더들이 생성형 AI 개인정보 보호 리스크를 줄이기 위해 가장 중요하게 여기는 분야

분산된 위치 전반에 걸친 보안 관리 통제



69%

거대언어모델(LLM) 사용에 대한 통제된 접근 및 모니터링



55%

인재 및 전문 역량 개발



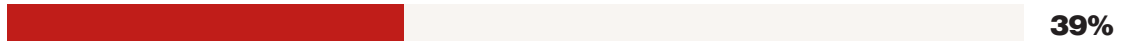
53%

거버넌스 프레임워크



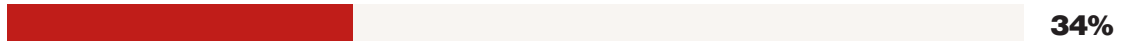
50%

퍼블릭 클라우드 데이터 노출



39%

다크 사이트 운영 지원



34%

*IT 리더들이 생성형 AI 개인정보 보호 리스크를 줄이는 데 가장 중요하다고 생각하는 분야는 무엇입니까? (상위 3개 항목 선정)

출처: FOUNDRY

생성형 AI가 여러 팀, 시스템 및 환경 전반으로 확장됨에 따라 구조적 과제는 더욱 증가할 것입니다.

이러한 복잡성을 방치할 경우, 규제 리스크, 예측 불가능한 운영 비용, 서비스 불안정, 그리고 더 근본적으로는 AI 기반 성과에 대한 경영진의 신뢰 저하와 같은 비즈니스 리스크로 바로 이어질 것입니다.

Nutanix의 코티켈라는 다음과 같이 말합니다. "AI의 변화 속도는 매우 빠르기 때문에, 이를 따라가지 못하면 사실상 경쟁

우위를 잃게 될 것이며, 데이터와 인프라는 비즈니스의 전략적 자산이며, AI 도입 과정에서 이를 잘못 다룰 경우 그 파급 효과는 훨씬 더 증폭될 것입니다."

IT 리더들은 초기 단계의 낙관론과 실제 운영 단계에서의 성공, 그리고 확장 가능한 기업 수준의 준비 태세 사이에서 균형을 잡아야 합니다. 생성형 AI를 규모 있게 확장하고 기대하는 혜택을 확보하기 위해서는 장기적인 성공에 필수적인 기반 요소들에 다시금 집중해야 할 것입니다.

단일 목적용 솔루션에서 플랫폼으로

파편화된 AI 솔루션은 비용, 거버넌스, 운영 회복 탄력성 측면에서 경영진의 노출 리스크를 증가시킵니다. 리스크를 줄이고 성공을 위한 신뢰할 수 있는 경로를 확보하기 위해, 최고 정보 책임자는 AI 시스템이 전략적으로 배포되고 회복 탄력성, 운영 단계, 보안을 갖추어 실행되도록 보장하면서 다른 모든 엔터프라이즈 애플리케이션과 동일한 방식으로 AI에 접근해야 합니다.

리스크를 줄이고 회복 탄력성을 높이는 최선의 방법은 일관된 거버넌스, 예측 가능한 비용, 회복 탄력성을 제공하는 단일 엔터프라이즈 AI 운영 모델을 확립하는 것입니다. 이러한 플랫폼 접근 방식은 도구를 단순히 통합하는 것이 아니라, 기존 IT 팀이 일관된 워크플로와 익숙한 운영 모델로 생성형 AI를 배포, 관리 및 운영할 수 있도록 AI 운영을 단순화하는 것입니다.

플랫폼이 없다면, 생성형 AI 이니셔티브는 비용 변동성을 높이고 책임 소재를 약화시키며 예측 가능한 비즈니스 성과를 제공하는 능력을 제한하는 고립된 단일 목적용 솔루션으로 파편화될 리스크가 있습니다.

주요 고려 사항은 다음과 같습니다:

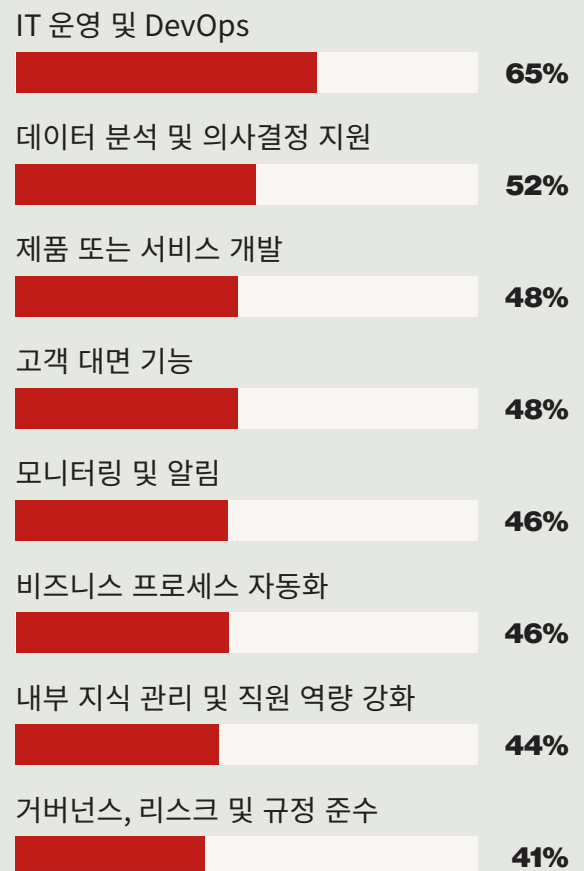
- **하이브리드 환경 전반의 지원:** AI 워크로드를 인프라를 한계치까지 밀어붙이고 있습니다. 성능, 안정성, 회복 탄력성은 AI를 지속 가능한 비즈니스 가치의 원천으로 만드는 데 필수적입니다. 그러나 AI의 연료가 되는 기업 데이터는 온프레미스 데이터 센터, 퍼블릭 클라우드, 엣지 등 여러 위치에 분산되어 있기 때문에, 최고 정보 책임자들은

에이전트형 AI, 기업의 준비 태세를 요구하다

생성형 AI 에이전트는 고객 응대용 챗봇을 넘어 핵심 운영 영역으로 이동하고 있습니다. 에이전트형 AI는 엔터프라이즈 AI 도입이 다음 단계로 넘어감을 의미하며, 복잡성 측면에서 일대 변화를 예고합니다. 단일 목적의 모델과 달리, 에이전트 시스템은 여러 AI 구성 요소가 자율적으로 상호 작용하며 데이터를 전달하고, 작업을 실행하며, 제한된 인간 개입만으로 의사결정을 내립니다.

그림 4

최고 정보 책임자들은 향후 12개월 동안 비즈니스 전반에 생성형 AI 에이전트를 배포할 계획입니다



출처: FOUNDRY

데이터, 지연 시간, 주권 요구 사항이 지시하는 곳 어디에서든 AI 워크로드를 실행할 수 있는 하이브리드 아키텍처를 배포해야 할 것입니다.

- **환경 전반에 걸친 일관된 거버넌스 (내장된 리스크 관리와 향상된 가시성, 책임성, 정책 시행 포함):** AI 모델링에 사용되는 민감한 데이터는 종종 국가 또는 조직의 경계 내에 유지되어야 하므로, AI 플랫폼은 데이터 레지던시를 준수하고 지역적 주권을 존중하며 데이터가 어디서 어떻게 처리되는지 제어할 수 있어야 합니다. 엔터프라이즈 환경에서 거대언어모델(LLM)을 배포하려면 강력한 검증 파이프라인, 지속적인 성능 모니터링, 그리고 리스크 및 규정 준수 정책과의 정렬이 필요합니다. AI 거버넌스 정책은 데이터 계보, 접근 제어, 사용 모니터링 및 수명 주기 정책 시행을 다루어야 할 것입니다.

← 7 페이지 사이드바에서 이어지는 내용입니다

이러한 에이전트형 AI 배포의 확대는 기업의 리스크 요소를 근본적으로 변화시킵니다. 각 에이전트는 어떤 데이터에 접근할 수 있고, 어떤 시스템에 영향을 미치며, 그 행동이 어떻게 모니터링되고 제한되어야 하는지 등 새로운 거버넌스 요구 사항을 발생시킵니다. 에이전트들이 연쇄적인 워크플로 안에서 함께 작업하기 시작하면, 데이터 노출, 비용 예측 가능성, 성능 및 책임성 전반에 걸쳐 리스크가 복합적으로 커집니다.

많은 기업이 자동화와 의사결정을 가속하기 위해 엔터프라이즈 AI 이니셔티브를 탐색하고 싶어 합니다. 하지만 성숙한 운영 모델 없이 에이전트를 배포하는 것은 흔히 통제의 파편화와 사각지대로 이어집니다. 고정된 애플리케이션을 위해 구축된 기존의 가드레일은 AI 시스템이 동적으로 진화하고 상호 작용할 때는 충분하지 않습니다.

최고 정보 책임자들에게 엔터프라이즈 AI는 중요한 교훈을 다시금 확인시켜 줍니다: 바로 미래의 혁신은 오늘의 기반에 달려 있다는 점입니다. 거버넌스, 가시성, 그리고 표준화된 AI 운영에 일찍 투자하는 기업들은 에이전트를 안전하고 자신 있게 도입할 수 있는 위치에 설 것입니다. 반면 그렇지 못한 기업들은 혁신 속도가 통제 범위를 넘어서게 되어, 창출되는 가치보다 더 빠르게 커지는 리스크에 직면하게 될지도 모릅니다.

■ 생성형 AI 투자에 대한 예측 가능한 비용 및 투자 대비 수익(ROI) 입증 능력:

파편화된 포인트 솔루션은 비용 변동성을 높이고 긍정적인 비즈니스 성과를 창출하는 능력을 제한합니다.

이에 비해, 엔드 투 엔드(end-to-end) 가시성을 갖춘 플랫폼은 성능과 비용 최적화를 모두 단순화하여, AI가 성장함에 따라 병목 현상을 식별하고 워크로드를 조정하며 비용을 통제하기 쉽게 만듭니다. 하이브리드 환경 전반에 걸친 생성형 AI 애플리케이션의 자동화된 운영 환경은 효율성을 개선하고 투자 대비 수익(ROI)을 달성하는 데 도움을 줍니다.

Nutanix의 코티켈라는 다음과 같이 말합니다. "AI 운영 환경은 물리적 인프라 계층부터 오케스트레이션, 플랫폼 서비스, 중앙 집중식 관리에 이르기까지 엔터프라이즈 AI를 운영하는 데 필요한 전체 스택을 통합합니다." "목표는 스택의 모든 수준에서 최상의 역량을 제공하는 플랫폼을 구축하는 것입니다."

생성형 AI 시대의 혁신 조력자로 거듭나기

생성형 AI 도입이 서비스형 소프트웨어(SaaS) 애플리케이션 및 엔터프라이즈 워크플로를 포함한 다양한 미션 크리티컬 프로덕션 환경으로 빠르게 이동함에 따라, 이는 혁신 이니셔티브에서 비용 구조, 회복

탄력성, 거버넌스 및 장기적 경쟁력에 직접적인 영향을 미치는 기업 운영 전략으로 전환되고 있습니다. 이것이 바로 최고 정보 책임자들이 핵심 시스템에 적용하는 것과 동일한 규율을 생성형 AI 환경에도 적용하되, 더 큰 규모와 속도로 수행해야 하는 이유입니다.

생성형 AI를 위한 적절한 사용 사례를 정의하는 것은 프로세스와 워크플로를 자동화하고 최적화하도록 설계된 운영 모델을 구상하는 데 중요합니다. 운영 모델은 AI 워크로드와 애플리케이션을 지원하기 위해 데이터 파이프라인과 인프라를 업데이트하는 청사진 역할을 합니다.

플랫폼 접근 방식은 최고 정보 책임자들이 생성형 AI를 시범 운영 단계에서 전사적 운영 단계로 전환하도록 돕고, 혁신을 가속화하며 긍정적인 비즈니스 성과를 창출하는 턴키 방식의 회복 탄력적이고 안전하며 확장 가능한 환경을 제공합니다. 이는 모든 최고 정보 책임자에게 성공적인 결과입니다.

생성형 AI에 대한 투자는 증가하고 있지만, 투자 회수는 어떠할까요?

IDC,¹에 따르면, 생성형 AI에 대한 핵심 IT 지출은 2024년 401억 달러에서 2027년 1,511억 달러로 급격히 증가할 것으로 예상되며, 이 기간을 전 세계 기업들에게 매우 중요한 AI 구축 단계로 정의하고 있습니다.

Foundry의 설문조사에 따르면, 대부분의 IT 리더들은 향후 2년간 생성형 AI 예산을 평균 12% 늘릴 계획입니다.

또한, 기업 리더들은 이러한 투자로부터 비교적 빠른 투자 회수를 기대하고 있습니다. 10명 중 8명 이상(84%)이 향후 12개월 내에 긍정적인 투자 대비 수익(ROI)을 달성할 수 있을 것이라고 보통 수준 이상의 자신감을 보였습니다. 거의 모든 응답자(97%)는 향후 2년 안에 생성형 AI 이니셔티브가 측정 가능한 비즈니스 가치를 제공할 것이라고 확신하고 있습니다 (그림 5 참조).

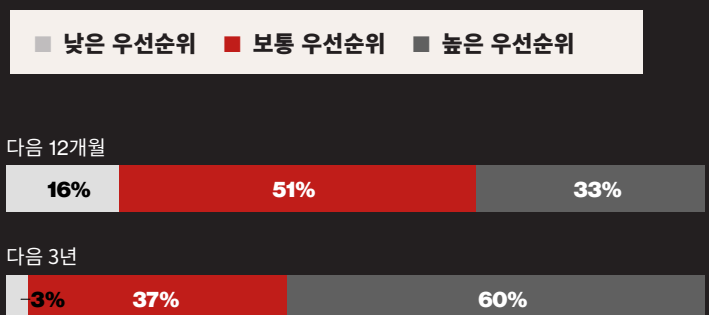
향후 3년 동안 생성형 AI의 성공적인 대규모 도입을 가능하게 할 핵심 역량이 무엇인지 묻는 질문에, 최고 정보 책임자와 기업 리더들은 다음을 꼽았습니다:

1. 생성형 AI 투자에 대한 예측 가능한 투자 대비 수익(ROI) 입증 능력
2. 혁신 및 실험 속도
3. 숙련된 인재 확보
4. 인프라의 회복 탄력성 및 성능
5. 거버넌스 및 책임 있는 생성형 AI 실천

투자 대비 수익(ROI)에 대한 압박이 거세짐에 따라, 최고 정보 책임자들은 생성형 AI 워크플로와 애플리케이션을 위해 단일 목적용 솔루션에서 플랫폼 중심의 접근 방식으로 전환해야 할 것입니다. 확장 가능한 플랫폼은 IT 리더들이 자율적인 워크플로, 데이터 주권, 그리고 엔드 투 엔드(end-to-end) 가시성을 제공할 수 있게 합니다. 이러한 생성형 AI 배포의 핵심 요소들은 시의적절하고 지속 가능한 투자 회수의 가능성을 높여줄 것입니다.

그림 5

IT 리더들의 생성형 AI 투자 대비 수익(ROI)에 대한 높은 기대치



출처: FOUNDRY

1 IDC, "생성형 AI 기회로 시장 내 성장을 창출하십시오."

엔터프라이즈 생성형 AI의 책임자는 누구인가? 모호함이 가져오는 숨겨진 리스크

생성형 AI가 기업 전반으로 확장되면서 많은 기업 내부에서 조용히 수면 위로 떠오르는 문제가 있습니다. 바로 불분명한 책임 소재(소유권)입니다. 조사 데이터에 따르면, 생성형 AI 거버넌스에 대한 책임은 의도적이든 혹은 어쩔 수 없는 상황에 의해서든 IT, 보안, AI 또는 데이터 담당 임원들 사이에 분산되어 있는 경우가 많습니다. 이러한 공유 모델은 AI가 가진 교차 기능적 특성을 반영하지만, 동시에 리스크를 초래하기도 합니다.

기존 애플리케이션과 달리 생성형 AI 시스템은 데이터 파이프라인, 모델, 인프라, 그리고 자동화된 의사결정을 모두 포괄합니다. 명확한 책임 소재가 없다면, 기업은 보안, 규정 준수, 비용 통제, 성능에 관한 일관된 표준을 시행하는 데 어려움을 겪게 됩니다. 이러한 공백은 팀의 의지가 부족해서가 아니라, 엔드 투 엔드(end-to-end) 운영 모델을 총괄하는 단일 임원이 없기 때문에 발생합니다.

최고 정보 책임자들에게 이러한 모호함은 도전인 동시에 기회입니다. 생성형 AI가 기업 인프라의 핵심 요소가 됨에 따라, 리더십의 기대치는 중앙 집중식 거버넌스와 분산식 실행의 결합으로 이동하고 있습니다. IT 리더들이 모든 AI 이니셔티브를 직접 통제할 필요는 없지만, 조직 전반에서 AI를 안전하게 배포하고, 관리하고, 모니터링하며 확장하는 운영 프레임워크를 정의하는 역할을 수행할 것으로 점점 더 많이 기대받고 있습니다.

생성형 AI의 책임 소재를 조기에 명확히 하는 기업은 AI가 미션 크리티컬한 워크플로로 확장됨에 따라 더 큰 통제력을 확보할 뿐만 아니라, 속도, 자신감, 그리고 신뢰성까지 얻게 됩니다.

생성형 AI 운영화를 위한 전략적 프레임워크

수십 년 동안 기업 IT는 도입과 운영 단계를 구분해 왔습니다. 하지만 생성형 AI는 이러한 경계를 허물고 있습니다. 많은 조직이 모델 출시와 시범 운영 프로젝트에 집중하지만, 장기적인 성공은 AI 시스템이 여러 환경에서 실제 배포된 이후에 어떻게 작동하느냐에 달려 있습니다.

생성형 AI는 기존의 애플리케이션 관리 방식을 뛰어넘는 새로운 운영 요구사항을 도입합니다. 최고 정보 책임자는 지속적인 모델 모니터링, 데이터 계보 추적, 에이전트 행동, 그리고 토큰 기반의 비용 변동성을 고려해야 합니다. 가시성은 단순한 가동 시간을 넘어 성능, 정확도, 사용 패턴, 정책 준수까지 확장되어야 합니다. 보안 또한 경계 방어를 넘어 모델 접근 제어 및 기계 간 신뢰 모델로 확장됩니다.

생성형 AI를 위한 운영 단계에는 자동화가 필수적입니다. 이전 워크로드에서 효과적이었던 수동 개입 방식은 AI 시스템이 증식하고 진화함에 따라 금세 지속 불가능해집니다. 기업은 하이브리드 환경 전반에 걸쳐 내장된 회복 탄력성, 자동화된 수정, 그리고 일관된 거버넌스가 필요합니다.

최고 정보 책임자가 정의하는 엔터프라이즈급 AI의 핵심 요소

- 예측 가능한 경제성
- 통제된 데이터 흐름
- 하이브리드 실행
- 운영 단계 자동화
- 비즈니스에 부합하는 성과

최고 정보 책임자들이 얻어야 할 교훈은 분명합니다: 생성형 AI를 특별한 사례의 워크로드로 취급해서는 안 된다는 것입니다. 생성형 AI는 전사적 자원 관리나 고객 관계 관리와 같은 엄격한 수준으로 운영되어야 하며, AI의 규모, 속도, 자율성에 적합한 새로운 통제 수단이 더해져야 합니다. 운영 단계를 마스터하는 기업만이 AI 목표를 지속 가능한 경쟁 우위로 바꿀 수 있을 것입니다.

이 주제에 대한 더 많은 경영진 인사이트와
사고 리더십을 보려면 [Nutanix Executive
Focus](#) 를 방문하십시오.

연구 조사 소개

본 리서치는 Nutanix의 후원을 받아 Foundry에서 실시한 온라인 설문조사 결과를 바탕으로 합니다. 해당 조사는 아시아 태평양, 북미, 서유럽 지역의 최고 정보 책임자 및 IT 고위 리더들을 대상으로 진행되었습니다. 응답자들은 직원 수 1,000명 이상의 기업에 재직 중이며, 소속 기업에서 생성형 AI 솔루션을 시범 운영했거나 이미 도입한 경험이 있습니다. 설문조사는 2025년 12월 7일부터 12월 23일까지 수행되었습니다.

Nutanix 소개

Nutanix는 하이브리드 멀티클라우드 컴퓨팅 분야의 선두주자로서, 전 세계 기업들이 어디에서나 애플리케이션을 실행하고, 엔터프라이즈 AI 워크로드를 배포하며, 데이터를 관리할 수 있도록 지원하는 통합 소프트웨어 플랫폼을 제공합니다. Nutanix와 함께라면 기업들은 기존 애플리케이션은 물론 최신 애플리케이션 운영을 단순화하여 비즈니스 목표에 더욱 집중할 수 있습니다. 전 세계 30,000개 이상의 고객사가 신뢰하는 Nutanix는 기업이 디지털 전환을 가속화하고, 하이브리드 멀티클라우드 환경을 일관성 있고 간편하며 비용 효율적으로 운영할 수 있도록 돕습니다. 자세한 정보는 www.nutanix.com에서 확인하실 수 있습니다.

CIO

후원:

NUTANIX