



Nutanix 엔터프라이즈 AI

Nutanix 엔터프라이즈 AI(NAI) 솔루션은 NVIDIA NIM 및 Hugging Face를 비롯한 주요 LLM 제공업체를 위한 엔드포인트 API를 제공하여 조직이 온프레미스 또는 퍼블릭 클라우드에서 다양한 GenAI 모델을 안전하게 배포할 수 있도록 지원합니다. NAI에는 간단한 UI, 역할 기반 액세스 제어(RBAC), 독립된 운영(다크 사이트 또는 에어 갭 배포) 기능이 포함되어 있어 엔터프라이즈 복원력, Day2 운영 및 규정 준수를 통해 AI 모델(LLM)의 운영, 모니터링 및 적응을 단순화할 수 있습니다.

주요 혜택

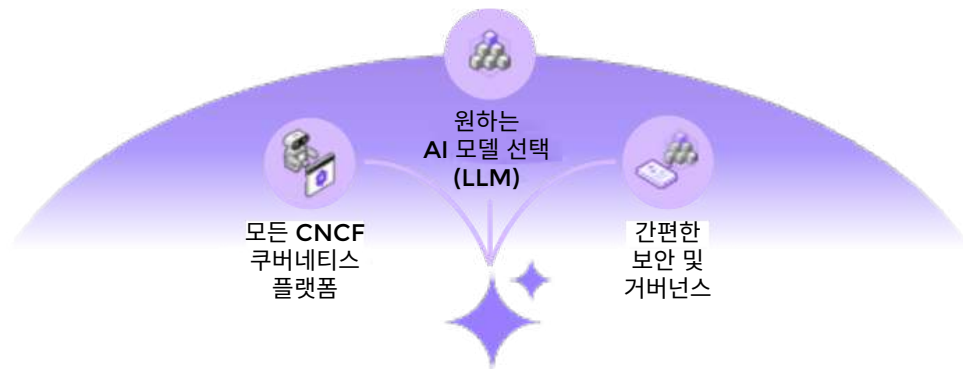
- 인공지능 도입을 간소화하고 LLM 배포를 단순화하며 API 워크플로우를 안전하게 보호하여 학습 곡선을 줄입니다.
- 사용자 친화적인 메트릭, 대시보드, 엔터프라이즈급 API 제공을 통해 확장 가능한 GenAI 앱 개발 지원합니다.
- LLM 및 API 배포, 모니터링, 보안을 위한 직관적인 인터페이스로 AI 운영을 단순화합니다.
- 원활한 통합, 검증된 LLM, 맞춤형 모델 옵션(비공개 업로드 포함)을 통해 GenAI 개발을 가속화합니다.
- Nutanix 쿠버네티스 플랫폼, 런처, 도커, Amazon EKS, Azure AKS, Google 클라우드의 GKE 등 모든 CNCF 인증 Kubernetes® 환경에 대한 지원으로 유연성을 확보하십시오.

IT 리소스를 AI 리소스로 만들기

Nutanix 엔터프라이즈 AI는 IT 리소스를 AI 리소스로 전환하여 워크플로우를 간소화하므로 생성형 AI LLM(대규모 언어 모델)을 신속하게 검색, 배포, 관리 및 테스트할 수 있습니다. 보안 API가 포함되어 있어 주요 엔터프라이즈 AI 앱이 원활하게 작동할 수 있습니다. 주요 메트릭과 모니터링을 통해 조직의 AI 역량을 강화하는 과정에서 발생하는 격차나 문제에 대한 가시성을 확보할 수 있습니다.

간단하면서도 유려한 사용자 인터페이스를 통해 LLM과 API를 쉽게 만들고, 관리하고, 보호할 수 있으며, 역할 기반 액세스 제어를 통해 팀 전체에서 안전하고 효율적인 확장을 지원합니다. 이 조합을 통해 단순화된 엔터프라이즈 AI 환경을 구축할 수 있습니다. 또한 Nutanix 엔터프라이즈 AI는 NVIDIA NIM 및 Hugging Face와 즉시 통합되어 필요한 검증된 LLM을 배포할 수 있습니다. 자체 모델을 업로드하여 전체 LLM을 제어할 수도 있습니다.

Nutanix 엔터프라이즈 AI를 사용하면 변화하는 GenAI 환경에 빠르게 적응하여 조직이 AI 모델, Kubernetes 아키텍처 및 인프라의 최신 개선 사항을 구현할 수 있는 기회를 극대화할 수 있습니다. Nutanix 쿠버네티스 플랫폼, 소프트웨어 파트너(런처, 도커 등), 베어메탈 또는 하이퍼스케일러 컨테이너 인스턴스(현재 AWS, Azure, Google Cloud)를 포함한 모든 CNCF 인증 Kubernetes 환경에 배포할 수 있습니다.



Nutanix 엔터프라이즈 AI

단순성

GenAI 애플리케이션 실행을 위한 AI 모델 선택 및 배포를 위한 깔끔하고 간편한 인터페이스입니다.

제어

IT 관리자가 애플리케이션 액세스, 엔드포인트 및 AI 모델을 제어할 수 있도록 AI 운영이 지원됩니다.

비용

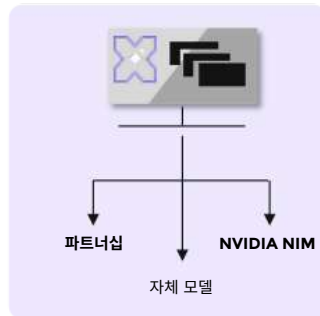
온프레미스 또는 퍼블릭 클라우드에서 AI를 실행하여 예측 가능한 총소유비용 (TCO)을 보장합니다.

1.



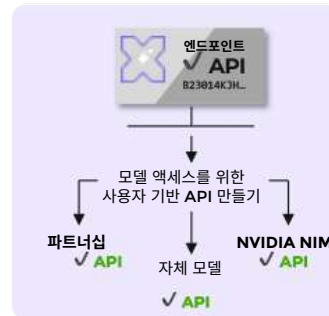
CNCF 인증 쿠버네티스에 Nutanix 엔터프라이즈 AI를 배포하고 실행합니다.

2.



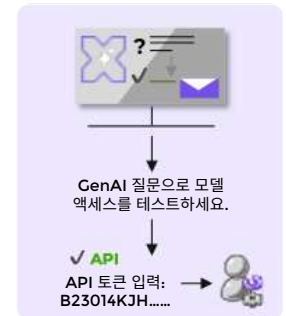
인터페이스에 로그인하여 간단히 LLM을 선택하고 배포하세요.

3.



보안 엔드포인트를 만들고 모델 액세스를 테스트합니다.

4.



개발자 또는 애플리케이션 소유자에게 자격 증명을 보냅니다.

API, 엔드포인트, LLM, 인프라
감사 및 모니터링

주요 기능:

LLM Value

- 유려한 엔터프라이즈급 사용자 인터페이스
- Hugging Face 또는 NVIDIA NIM의 AI 모델(LLM) 선택
- 자체 AI 모델(LLM) 업로드하기

AI 엔드포인트 관리

- API 토큰 생성 및 관리
- Hugging Face 또는 NVIDIA NIM의 파트너 API 토큰 관리
- API 코드 샘플

보안 및 감사 관리

- 역할 기반 액세스 제어(RBAC)
- AI 모델 프리플라이트 테스트
- AI 모델 및 API 모니터링
- Kubernetes 리소스 모니터링
- GPU 사용량 모니터링 이벤트 감사
- 통합 Nutanix 펄스 보고

리소스 및 시작 방법

엔터프라이즈 AI가 기업 기술 스택의 핵심 요소로 자리 잡으면서, IT 리소스를 AI 리소스로 전환하고, 간단한 AI Day 2 운영(LLMOps)을 가능하게 하며, 생성형 AI(GenAI)의 다음 큰 진보에 유연하게 대응할 수 있는 능력은 더 이상 단순히 바람직한 것이 아니라 필수적입니다.

Nutanix 엔터프라이즈 AI에 대해 자세히 알아보려면 <https://www.nutanix.com/products/nutanix-enterprise-ai>를 방문하거나 [nutanix.com/kr/testdrive](https://www.nutanix.com/kr/testdrive)에서 테스트 드라이브를 체험해 보세요.

Nutanix 엔터프라이즈 AI 설정 시작하기: https://portal.nutanix.com/page/documents/details?targetId=Nutanix-Enterprise-AI-v1_0:Nutanix-Enterprise-AI-v1_0

Nutanix 엔터프라이즈 AI 라이선스 방법 알아보기: <https://www.nutanix.com/products/cloud-platform/software-options#nai>

자세한 내용은 www.nutanix.com/kr을 참조하십시오.

NUTANIX

info@nutanix.com | www.nutanix.com/kr | [@nutanix](https://twitter.com/nutanix)

©2024 Nutanix, Inc. All rights reserved. Nutanix, Nutanix 로고 및 여기에 언급된 모든 Nutanix 제품 및 서비스 이름은 미국 및 기타 국가에서 Nutanix, Inc.의 등록 상표 또는 상표입니다. 쿠버네티스는 미국 및 기타 국가에서 The Linux Foundation의 등록 상표입니다. 여기에 언급된 기타 모든 브랜드명은 구분을 위한 목적으로만 사용되었으며 각 해당 소유주(들)의 상표일 수 있습니다.