

Panoramica di GPT-in-a-Box

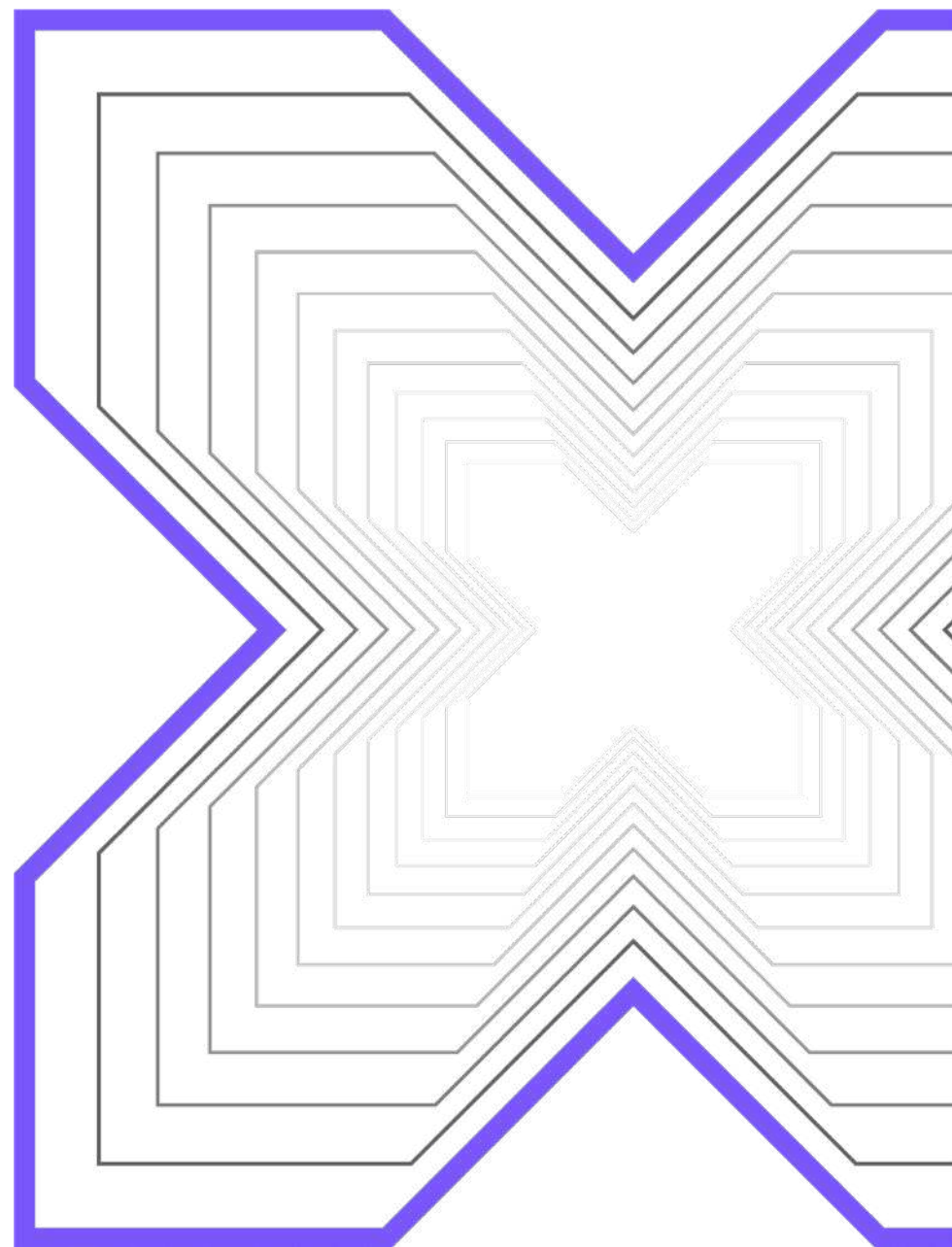
Nutanix GPT-in-a-Box™ è una soluzione AI pronta all'uso per le organizzazioni che desiderano implementare funzionalità GPT (Generative Pre-trained Transformer) mantenendo il controllo sui propri dati e sulle proprie applicazioni.

La soluzione risolve i problemi di complessità, scalabilità e sicurezza che le aziende si trovano ad affrontare quando adottano applicazioni di intelligenza artificiale generativa e AI/ML, grazie a una piattaforma full-stack software-defined pronta per l'intelligenza artificiale che semplifica e accelera le iniziative dal datacenter sino all'edge.

Nutanix GPT-in-a-Box fornisce gli strumenti necessari per semplificare la tua infrastruttura in modo che sia pronta all'implementazione dell'intelligenza artificiale, garantendo che i contenuti AI rimangano sotto il totale controllo aziendale in modo da soddisfare i requisiti di sicurezza, privacy e conformità. La soluzione permette di creare, perfezionare ed eseguire modelli di vario tipo (inclusi GPT e LLM) in modo efficiente sfruttando i team e le competenze esistenti.

Caratteristiche

- **Libertà di scelta:** il modello Bring-your-own-AI ti consente di scegliere tra i modelli più diffusi, inclusi Meta Llama 2, Falcon LLM e MosaicML MPT.
- **Operatività semplificata:** coniuga le esigenze di sviluppatori e data scientist grazie alle funzionalità di governance e al provisioning self-service.
- **Servizi dati enterprise:** proteggi i modelli di intelligenza artificiale grazie al valore comprovato dei servizi dati Nutanix.
- **Riduzione del TCO:** alloca le risorse in modo dinamico su tutta la tua infrastruttura — dal cloud al core fino all'edge — e crea valore in maniera immediata ed efficiente.



GPT-in-a-Box in breve

Software Nutanix Cloud Platform	Servizi di pianificazione, progettazione e implementazione*	Stack MLOps
<p>Sono inclusi:</p> <ul style="list-style-type: none">Nutanix Cloud Infrastructure UltimateNutanix Unified Storage Pro con 10 TB di storage per oggetti e file <p>Servizi di dati completi, virtualizzazione integrata con l'hypervisor AHV®, resilienza e funzionalità di gestione grazie all'AI Ops per l'infrastruttura: monitoraggio, pianificazione, dimensionamento e automazione low-code.</p> <p>Visita la nostra pagina dedicata alle opzioni software per un elenco completo delle funzionalità incluse nella Nutanix Cloud Infrastructure Ultimate e in Nutanix Unified Storage Pro.</p>	<p>Sono inclusi:</p> <ul style="list-style-type: none">Workshop di pianificazione: il Nutanix Artificial Intelligence/Machine Learning Planning Workshop offre ai team IT una guida pratica e approfondita per pianificare e implementare carichi di lavoro AI/ML basati su un modello GPT per i cluster Nutanix Cloud Infrastructure™ on-premise. Questo workshop è particolarmente utile durante la fase di pianificazione della soluzione Nutanix GPT-in-a-Box.Workshop di progettazione dello stack: il Nutanix Artificial Intelligence/Machine Learning Stack Design Workshop offre ai team IT una guida pratica e approfondita per creare un progetto completo per i carichi di lavoro AI/ML sulla base di un modello GPT per i cluster Nutanix Cloud Infrastructure™ on-premise. Questo workshop, incentrato su aspetti quali prestazioni, scalabilità, flessibilità, integrazione ed esigenze operative, è particolarmente utile durante la fase di progettazione della soluzione Nutanix GPT-in-a-Box.Implementazione: il servizio Nutanix Artificial Intelligence Machine Learning Deployment accelera l'implementazione di una soluzione AI/ML basata su un carico di lavoro GPT sulla Nutanix Cloud Platform™. Un consulente altamente qualificato implementa le istanze di calcolo AI e i prerequisiti associati per utilizzare il cluster Nutanix Cloud Infrastructure™ per l'inferenza. Questa offerta è ideale per la fase di implementazione della soluzione Nutanix GPT-in-a-Box.	<p>Sono inclusi:</p> <ul style="list-style-type: none">Software PyTorch®: un framework completo per la creazione di modelli di intelligenza artificiale generativa.Set di strumenti Kubeflow® MLOps: consente di distribuire flussi di lavoro ML su container Kubernetes® per garantire maggiore semplicità, portabilità e scalabilità.Modello Bring-your-own-AI: scegli tra i modelli open source più diffusi, inclusi Meta Llama 2™, Falcon LLM™ e MosaicML® MPT.

*I partner possono fornire servizi di implementazione e migrazione equivalenti al posto dei servizi offerti da Nutanix.



Servizi professionali per GPT-in-a-Box

Codice prodotto	Descrizione del servizio
CNS-DAY0-WRK-PLN-GPT/AI/ML	Workshop di pianificazione
CNS-INF-A-WRK-DES-GPT/AI/ML	Workshop di progettazione dello stack
CNS-INF-A-SVC-DEP-GPT/AI/ML	Implementazione

Per ulteriori informazioni, visita le [Descrizioni dei servizi nel nostro Catalogo dei Servizi Professionali](#).

Use case/carichi di lavoro suggeriti

Per tutti i settori:

- Bot AI per l'assistenza clienti e il supporto tecnico
- Ricerca e analisi di documenti quali per esempio prestiti, contratti, studi
- Co-piloting per il codice
- Protezione contro frodi e furti

Settore medico e farmaceutico: progettazione generativa di anticorpi, elaborazione di dati *-oma* per la scoperta di farmaci

Settore finanziario: ottimizzazione dei portafogli, analisi per l'upselling ai clienti

Settore retail/edge: ispezioni visive della qualità, rilevamento delle frodi

Settore federale: funzionalità copilot e simulazioni — digital twin, sistemi multi-agente

Per ulteriori informazioni visita il sito [Nutanix.com](https://www.nutanix.com) o contatta il [partner Nutanix più vicino](#).

NUTANIX

info-italy@nutanix.com | www.nutanix.com | [@NutanixItaly](https://twitter.com/NutanixItaly)

© 2024 Nutanix, Inc. Tutti i diritti riservati. Nutanix, il logo Nutanix e tutti i nomi di prodotti e servizi citati nel presente documento sono marchi o marchi registrati di Nutanix, Inc. negli Stati Uniti e in altri Paesi. Tutti gli altri nomi di marchi qui menzionati sono solo a scopo identificativo e potrebbero essere marchi commerciali di proprietà dei rispettivi titolari. GPT-Kit-Nutanix-HCI-Starter-Kit-Customer-Version-2-FY24Q1-it-IT-12022024