TheFuturum Group

Aumenta l'efficienza e ottimizza le prestazioni con soluzioni aziendali AI pronte all'uso

AUTORE

Steven Dickens
Chief Technology Advisor | The Futurum Group

Ron Westfall
Direttore della ricerca | The Futurum Group

AGOSTO 2024

IN PARTNERSHIP CON

Lenovo | NVIDIA

Riepilogo esecutivo

L'AI è diventata una priorità assoluta per le aziende e le organizzazioni, ma la sua rapida adozione deve affrontare sfide di varia natura. Nonostante la popolarità delle applicazioni di AI generativa pubblica, le organizzazioni IT devono affrontare problemi come la mancanza di conoscenze specifiche del dominio, una prontezza aziendale limitata, i problemi di sicurezza e le competenze insufficienti. Questi problemi sono ulteriormente aggravati da soluzioni di gestione dei dati inadeguate, normative di conformità e costi elevati relativi alle commissioni cloud e al consumo energetico. Il mercato frammentato dei fornitori di software indipendenti (ISV) per le applicazioni di AI basate rende la situazione ancora più complessa, rallentando potenzialmente i progressi.

L'AI ibrida emerge come soluzione strategica per le sfide da affrontare nell'adozione dell'AI. Questo approccio combina apprendimento automatico, deep learning e reti neurali con l'esperienza umana per creare modelli di AI altamente accurati e specifici del dominio. Utilizzando un mix di risorse di AI private e pubbliche, le aziende possono ottenere vantaggi da entrambi: il controllo e la personalizzazione dell'infrastruttura privata insieme alla scalabilità e alle funzionalità dell'AI basata sul cloud. Questo modello ibrido consente alle organizzazioni di ottimizzare la propria strategia di AI in base alle loro esigenze specifiche, al budget e ai requisiti tecnologici.

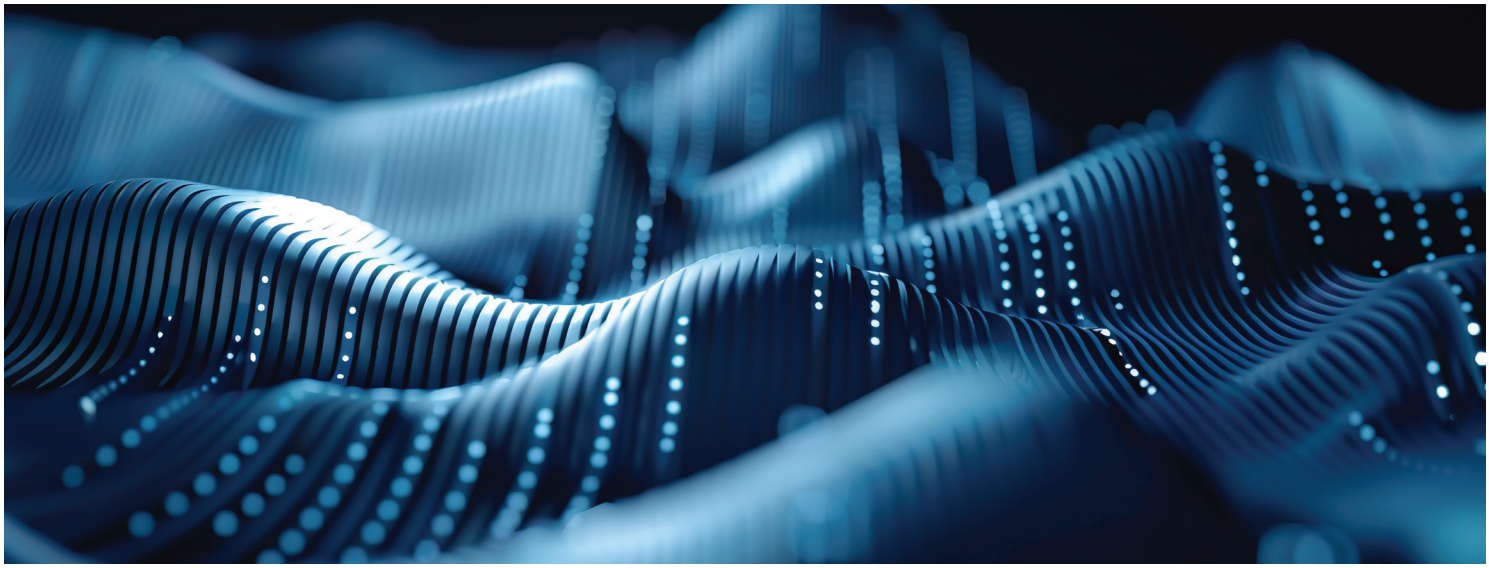
Per i clienti che hanno dubbi circa la privacy dei dati, la sicurezza o il controllo dei propri sistemi di AI, l'infrastruttura privata di AI si pone come una soluzione valida per loro. Ospitando modelli di AI ed elaborando sulla propria infrastruttura sicura, le aziende possono mantenere il pieno controllo e la visibilità su come vengono utilizzati i loro dati e sistemi di AI. Ciò può essere particolarmente importante per i settori con normative sui dati rigorose, proprietà intellettuale sensibile o applicazioni mission-critical. L'infrastruttura AI privata consente inoltre una maggiore personalizzazione e integrazione con i sistemi IT e i flussi di lavoro esistenti di un'azienda.

In questo scenario, i responsabili aziendali per le decisioni IT (ITBDM) come CIO, CTO e Head of Architecture, possono scoprire come l'AI privata consenta di personalizzare i modelli di AI alle esigenze specifiche della loro organizzazione con un maggiore controllo sull'intero stack di AI, inclusi modelli, applicazioni, infrastruttura e framework operativi. Questo migliore controllo consente alle aziende di soddisfare requisiti specifici circa personalizzazione, flessibilità, privacy, inferenza in tempo reale, conformità e sicurezza. Di conseguenza, l'AI ibrida con modelli privati diventa un potente strumento per le organizzazioni che cercano di sfruttare il potenziale dell'AI affrontando le complessità dell'implementazione e della gestione dei dati.

Questo documento illustra come le organizzazioni possono progettare e costruire soluzioni di AI privata e perché l'AI privata risponde alla crescente domanda di un maggiore controllo sulle soluzioni di AI, tra cui protezione e trasparenza. La strategia Lenovo Smarter AI for All dà la priorità all'AI aziendale, sfruttando la sua potenza per guidare la trasformazione intelligente in tutti i settori e consentire alle organizzazioni di innovare rapidamente e mantenere un vantaggio tecnologico rispetto alla concorrenza.

© 2024 The Futurum Group. Tutti i diritti riservati.

Scopri perché la creazione di soluzioni vincenti di AI privata richiede un ecosistema di partner, competenze in materia di AI, un gruppo attentamente curato di ISV e delle loro app di AI, un'infrastruttura di dimensioni adeguate per carichi di lavoro mirati e il supporto degli sviluppatori per garantire l'adozione e l'innovazione dell'AI a livello di organizzazione.



Sezione 1: Il panorama dell'AI privata

Con l'accelerazione dell'adozione, la protezione dei dati sensibili è diventata fondamentale. Le tecniche di AI privata, come la privacy differenziale e il calcolo crittografato, offrono una soluzione potente, consentendo alle organizzazioni di garantire un addestramento di successo del modello di AI mantenendo una solida sicurezza.

Con l'affermarsi dell'AI privata, possiamo identificare i settori che sono pronti a trarre il massimo vantaggio dall'AI privata oggi:

- Settori altamente regolamentati come i servizi finanziari, l'assistenza sanitaria, il settore pubblico, le utenze, il petrolio e gas
- Settori distribuiti con grande presenza dell'edge, sedi distribuite come vendita al dettaglio, ospedali, aeroporti e stadi
- Organizzazioni che hanno già importanti dati e data center in loco che supportano grandi ambienti privati
- Settori che dipendono da assistenti AI altamente addestrati/perfezionati e specifici per dominio, come organizzazioni in camoi quali diritto, ingegneria e medicina

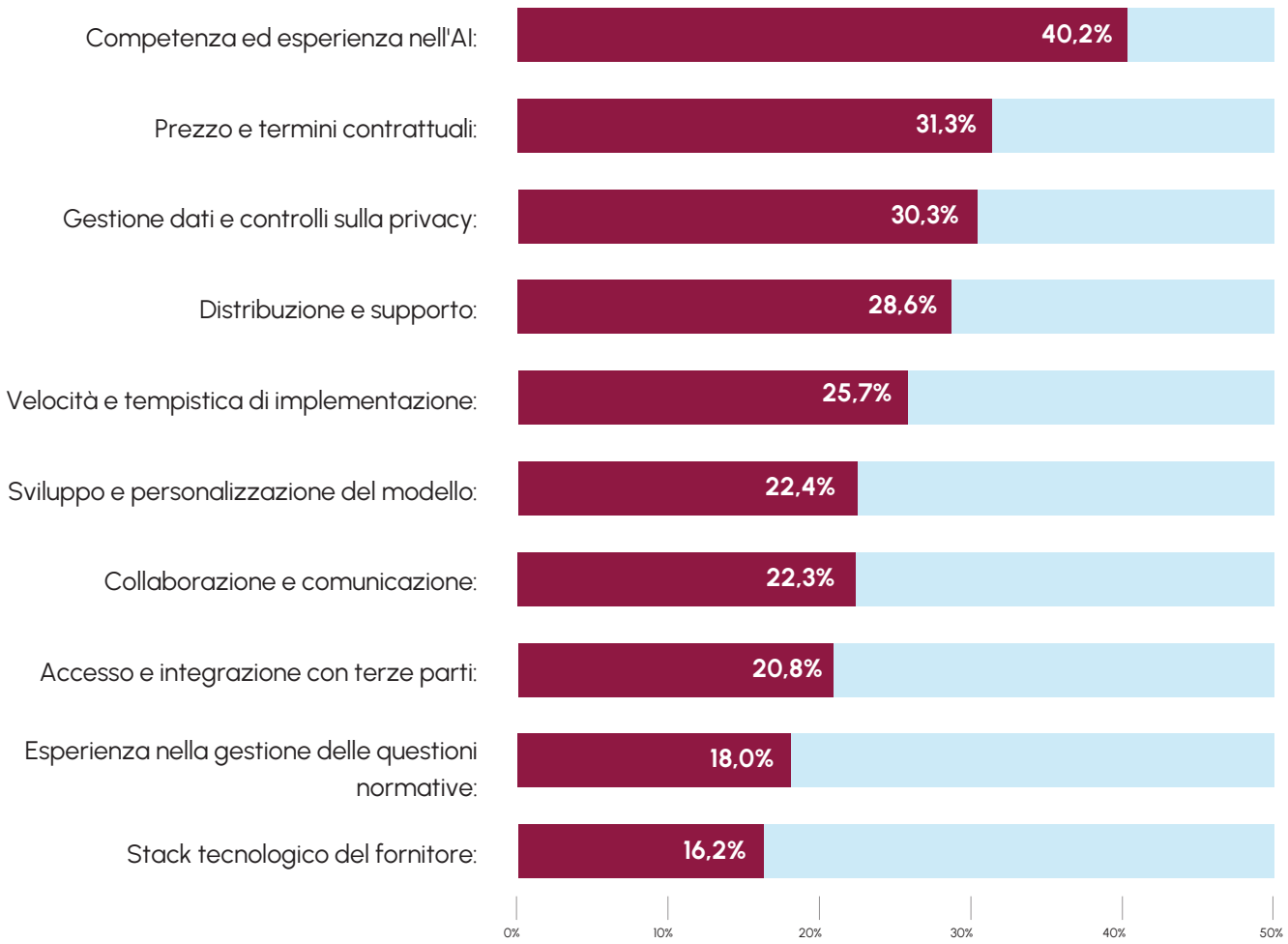
Nel panorama aziendale dinamico di oggi, le aziende (e gli sviluppatori) stanno adottando un approccio sofisticato all'implementazione dell'AI. Stanno combinando strategicamente vari elementi, da modelli piccoli e grandi a diversi set di dati, e dalle CPU alle GPU, per creare soluzioni di AI privata ottimizzate. Questa personalizzazione consente loro di bilanciare fattori cruciali come costo, sicurezza, conformità, prestazioni e latenza, rendendo l'AI privata una realtà pratica per le loro esigenze specifiche.

© 2024 The Futurum Group. Tutti i diritti riservati.

Inoltre, le organizzazioni che esplorano l'AI generativa (GenAI) stanno adottando un modello ibrido. Questo approccio integra modelli di GenAI sia pubblici che privati, sfruttando i punti di forza di ciascuno. I modelli pubblici offrono accessibilità e funzionalità predefinite per la generazione di testo, immagine, video e codice. Al contrario, i modelli privati, sia self-hosted che basati sul cloud, offrono una personalizzazione avanzata. Questi modelli privati sono spesso basati su modelli pubblici nelle fondamenta, incorporando applicazioni su misura che offrono output specializzati, controlli dei dati più forti e misure di sicurezza avanzate. Questa strategia ibrida consente alle aziende di sfruttare la potenza della GenAI mantenendo la flessibilità per soddisfare i loro requisiti e vincoli unici.

Scegliere il fornitore giusto

Nell'attuale panorama dell'AI, le organizzazioni stanno evolvendo i criteri di selezione dei fornitori. Quando scelgono i partner di AI, le aziende ora danno uguale importanza alle capacità dell'AI privata e alle competenze generali di AI. In particolare, danno la priorità ai fornitori che eccellono nella gestione sicura dei dati e implementano controlli sulla privacy robusti, dimostrando anche una vasta esperienza e competenza tecnica nello sviluppo e nella distribuzione dell'AI. Questo cambiamento riflette la crescente consapevolezza che le soluzioni di AI efficienti devono non solo essere innovative, ma anche affidabili e conformi agli standard di protezione dei dati:



Fonte: The Futurum Group

© 2024 The Futurum Group. Tutti i diritti riservati.

Le organizzazioni stanno rapidamente adottando soluzioni di AI privata, dando la priorità ai fornitori che soddisfano i loro criteri in continua evoluzione. Questo cambiamento è guidato dai vantaggi chiave inerenti all'AI privata:

1. Misure di sicurezza rafforzate.
2. Inferenza AI conveniente nel lungo termine.
3. Conformità alle normative costante.
4. Capacità di personalizzazione.
5. Migliori framework di governance.

L'AI privata è particolarmente preziosa per le organizzazioni che operano secondo severe leggi sulla privacy dei dati, come il GDPR europeo o il CCPA della California. Garantendo che i dati siano gestiti in modo sicuro e in conformità con queste normative, l'AI privata aiuta le aziende a navigare in scenari legali complessi sfruttando al contempo il potenziale dell'AI.

I regolamenti specifici per l'AI stanno per essere integrati, considerando che la legge sull'AI dell'UE stabilisce una struttura di governance a livello sia europeo che nazionale che richiede una valutazione di conformità prima di distribuire sistemi di AI ad alto rischio e prevede dei meccanismi per l'applicazione. Inoltre, Stati Uniti, Regno Unito e Cina stanno sviluppando normative simili. In linea di massima, un'organizzazione richiederà un modello privato se:

- **La privacy dei dati è fondamentale:** un operatore sanitario potrebbe utilizzare un modello GenAI per sviluppare un sistema di diagnosi del paziente addestrato sui dati dei pazienti. Ciò consentirebbe al fornitore di offrire ai pazienti diagnosi private più accurate e personalizzate.
- **Sono necessarie informazioni transazionali aggiornate:** un rivenditore potrebbe utilizzare un modello GenAI per sviluppare un sistema di prezzi dinamico che viene aggiornato con i dati di vendita più recenti. Ciò consentirebbe al rivenditore di ottimizzare i prezzi e massimizzare i profitti.
- **Vi è l'opportunità di sfruttare dati proprietari e proprietà intellettuale:** un fornitore di servizi finanziari potrebbe utilizzare un modello GenAI per sviluppare una strategia di trading di investimento proprietaria addestrata sui suoi dati e sulla sua ricerca.

Capacità di personalizzazione: GenAI privata

L'AI privata è diventata fondamentale per garantire la priorità assoluta delle aziende: adottare con successo la tecnologia GenAI. Questo cambiamento evidenzia la crescente importanza di sfruttare in modo efficiente i modelli di trasformatore generativo pre-addestrato (GPT) privato. In tal modo, le aziende possono creare i propri sistemi di GenAI sicuri e motori decisionali privati basati sull'AI. A nostro avviso, le organizzazioni che mirano a creare un cloud privato per l'AI dovrebbero concentrarsi su queste considerazioni chiave:

- Piattaforme integrate di livello aziendale in grado di ospitare governance, gestione dei dati e gestione ML generalmente costruite per le aziende Fortune 500.
- AI e stack di dati appositamente progettati per offrire una scelta di app, agenti e assistenti, competenze AI e completarli attraverso ISV e un ecosistema di partner convalidati.

© 2024 The Futurum Group. Tutti i diritti riservati.

- Selezione ed esplorazione della fusione di piattaforme open source e chiuse, come la piattaforma NVIDIA AI Enterprise, insieme ad altri proprietari di stack di piattaforme dati.

- Dai la priorità all'esame delle piattaforme open source che possono essere personalizzate per le applicazioni di AI.

- L'edge deve soddisfare i requisiti circa dati e latenza che sono vitali per creare soluzioni edge all'edge con gli ISV.

- Nello scegliere le funzionalità di generazione aumentata di recupero (RAG), disporre di una knowledge base ben definita migliora la definizione del dominio o del contesto specifico al fine di migliorare le risposte del modello linguistico.

L'ecosistema AI: opportunità e sfide

Il panorama dell'AI si sta evolvendo rapidamente, con nuovi fornitori di software indipendenti (ISV) e applicazioni di AI che emergono man mano che i modelli open source diventano più potenti e convenienti. Piattaforme come Hugging Face Hub, che ospita oltre 350.000 modelli, 75.000 set di dati e 150.000 app demo, sono un esempio di questa tendenza. Questo ecosistema in espansione consente alle aziende di personalizzare le soluzioni di AI privata in base alle loro esigenze specifiche.

Tuttavia, le aziende devono affrontare sfide importanti nella navigazione in questo panorama complesso e nel rimanere aggiornate con i nuovi sviluppi. Collaborare con degli esperti fidati può aiutare le organizzazioni a:

1. Convalidare e integrare diversi elementi dell'ecosistema.

2. Accedere a competenze critiche di AI e dati.

3. Selezionare l'infrastruttura AI ottimale per un equilibrio tra prezzo, prestazioni e sicurezza.

Uno sviluppo responsabile dell'AI

Uno sviluppo etico dell'AI si basa su una gestione e una protezione responsabili dei dati. Ciò garantisce che i modelli di AI siano addestrati su dati rappresentativi, imparziali e accurati. Le aziende danno la priorità alla flessibilità nella scelta di soluzioni di AI privata che proteggono e gestiscono al meglio i loro carichi di lavoro di AI.

I fattori chiave per l'AI pronta per l'azienda includono:

1. L'adesione ai principi di AI responsabile.

2. L'implementazione di pratiche MLOps solide.

3. Sviluppo di strategie complete di gestione dei dati in tutte le sedi organizzative.

Questi elementi sono fondamentali per creare fiducia nelle applicazioni e nei modelli di AI, rendendoli veramente pronti per l'azienda.

© 2024 The Futurum Group. Tutti i diritti riservati.

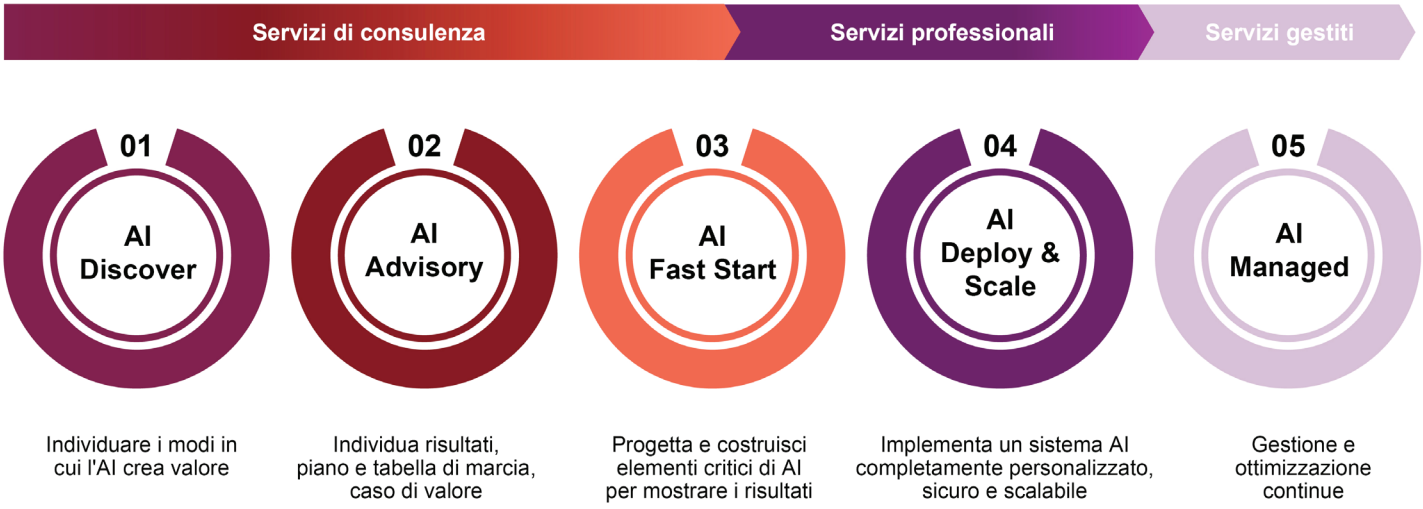
Sezione 2: Soluzioni di AI aziendale di Lenovo

Lenovo Smarter AI for All vision offre una vasta gamma di soluzioni di AI, dai dispositivi di AI personali all'infrastruttura AI di livello aziendale. Le soluzioni di AI aziendale Lenovo sono progettate per essere sicure, efficienti e scalabili, aiutando le aziende a utilizzare l'AI in modo responsabile e corretto.

Rileviamo che Lenovo offre una suite completa di soluzioni di AI aziendale che danno la priorità alla protezione e alla gestione dei dati. Queste soluzioni trovano un equilibrio tra flessibilità, sicurezza ed efficienza, consentendo alle organizzazioni di sfruttare il potenziale dell'AI salvaguardando le informazioni sensibili. L'approccio di Lenovo include funzionalità di AI sul dispositivo e on-premise, garantendo che i dati rimangano isolati dalle reti pubbliche. Questa strategia consente alle aziende di implementare l'AI in modo sicuro e responsabile, sia all'interno dei singoli dispositivi che sull'infrastruttura server in loco. Concentrandosi su questi aspetti critici, Lenovo consente alle aziende di sfruttare tecnologie di AI avanzate mantenendo al contempo una rigorosa sicurezza dei dati e la conformità normativa.

Le organizzazioni che intendono creare un'AI privata devono iniziare con la comprensione degli obiettivi aziendali e delle capacità dei dati. Sebbene questo processo importante sia multiforme e dipenda dalle capacità e dal focus dell'organizzazione, i seguenti sei passaggi possono portare al successo del progetto:

1. Individua il problema aziendale e il risultato desiderato.
2. Estrai e analizza i dati.
3. Costruisci il caso aziendale.
4. Progetta il modello operativo, l'infrastruttura, il design del modello e la distribuzione.
5. Costruisci/addestra o perfeziona il modello.
6. Distribuisci, monitora le prestazioni e affina secondo le necessità.



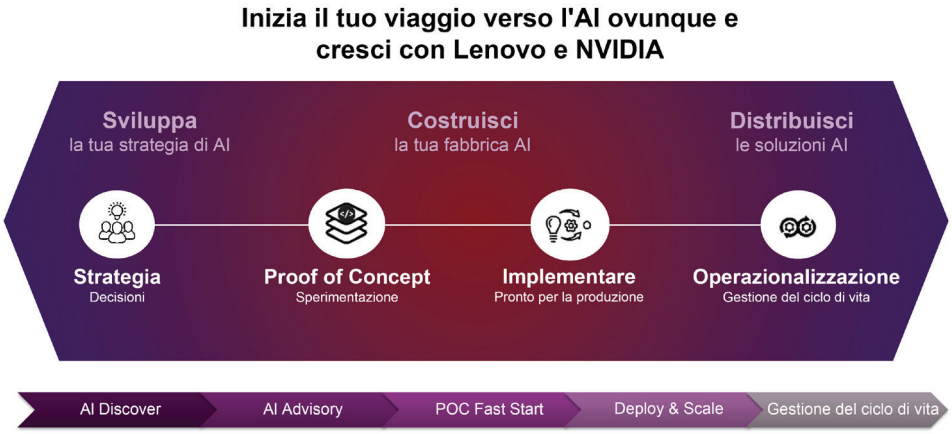
Fonte: Lenovo

© 2024 The Futurum Group. Tutti i diritti riservati.

I team Lenovo accelerano i processi attraverso il programma AI CoE e AI Innovators con partner e ISV. Le aree di soluzione chiave includono:

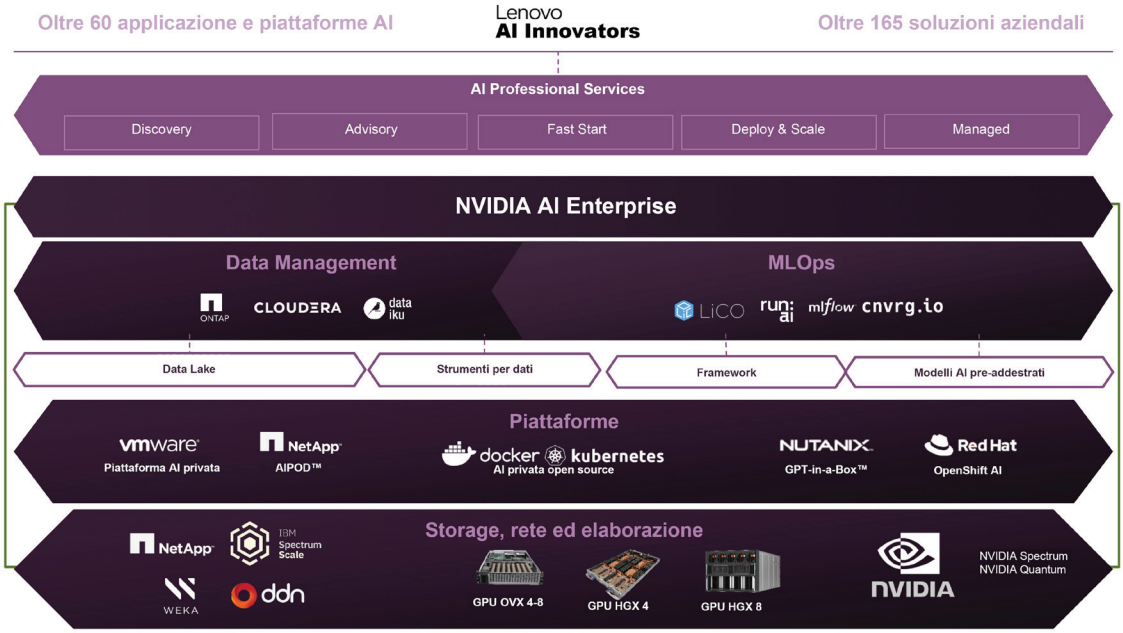
PC e workstation AI che supportano dipendenti, data scientist e sviluppatori

- Soluzioni di gestione e protezione dei dati
- Servizi e competenze AI in grado di accelerare i processi di progettazione, strategia ed esecuzione
- Soluzioni AI Fast Start che offrono casi d'uso e progetti pilota per il settore con i dati delle organizzazioni
- Soluzioni di infrastruttura AI ad alta efficienza energetica per esigenze di Edge e datacenter che offrono elaborazione dalle giuste dimensioni (CPU, GPU), storage, innovazioni di raffreddamento a liquido e scelta



Fonte: Lenovo

Lenovo sta creando un'architettura modulare che offre flessibilità in grado di adattarsi alle esigenze uniche dei clienti creando allo stesso tempo una standardizzazione mediante elementi costitutivi comuni. Questo parte dai casi d'uso e dalle esigenze aziendali e termina a livello di server e storage e questo concetto viene portato avanti attraverso l'architettura della soluzione di Lenovo per i clienti che vogliono costruire la loro fabbrica o centro di competenza di AI.



Fonte: Lenovo

© 2024 The Futurum Group. Tutti i diritti riservati.

Dal nostro punto di vista, spiccano il portafoglio di AI aziendale Lenovo e l'alleanza con NVIDIA. Utilizzando la piattaforma software NVIDIA AI Enterprise, che include il framework NVIDIA NeMo, il nuovo design di riferimento per GenAI basato su LLM mostra alle aziende come distribuire e commercializzare strumenti e modelli di base GenAI sfruttando una soluzione pre-convalidata, completamente integrata e ottimizzata per le prestazioni. Grazie alla partnership con NVIDIA e alla collaborazione per lo sviluppo delle soluzioni, Lenovo può aiutare le aziende a migliorare i flussi di lavoro LLM e a massimizzare prestazioni, sicurezza e dimensionamento.

Il panorama delle infrastrutture AI è complesso, con una notevole incertezza sui componenti necessari per le architetture RAG e i modelli linguistici di grandi dimensioni. Lenovo affronta questa sfida offrendo soluzioni per infrastrutture modulari e scalabili. Queste vanno dalle configurazioni standard alle opzioni ad alte prestazioni, consentendo ai clienti di prendere decisioni informate in base alle loro esigenze specifiche.

Le offerte chiave includono:

1. Server NVIDIA OVX con configurazioni a 4 o 8 GPU per inferenza e architetture RAG. Utilizza le GPU NVIDIA L40S, insieme alle tecnologie di rete NVIDIA Connect-X e BlueField aggiornabili a NVIDIA H100NVL o HGX 8-GPU per prestazioni migliorate.

2. Rete AI ad alte prestazioni con piattaforme NVIDIA Spectrum Ethernet o NVIDIA Quantum InfiniBand.

3. Soluzioni di storage flessibili che supportano l'infrastruttura definita da software (SDI). Opzioni Lenovo DM / NetApp. Storage ad alte prestazioni da Weka o DDN.

Questo approccio consente alle organizzazioni di personalizzare la propria infrastruttura AI, bilanciando i requisiti di prestazioni con i vincoli di budget e le esigenze di scalabilità future.

Lenovo adotta un approccio simile alle topologie e alle architetture di rete, fino allo stack software, andando dalle offerte open source a quelle aziendali completamente supportate (ad esempio NVIDIA AI Enterprise, NVIDIA NIM, OpenShift AI, ecc.). Inoltre, ove applicabile, l'azienda sfrutta i partner Lenovo AI Innovator per fornire il livello di applicazione determinato in base al caso d'uso. Inoltre, Lenovo ha sviluppato servizi professionali in grado di supportare i clienti lungo il loro percorso di adozione, dall'individuazione alla distribuzione e alla gestione.

Nell'era in cui i dati sono il nuovo "petrolio" e l'AI la raffineria, Lenovo ha sviluppato una strategia completa per massimizzare il valore per i suoi clienti. Questo approccio si estende dall'infrastruttura agli strumenti software, garantendo che i dati delle organizzazioni siano pronti per l'AI, governati correttamente e gestiti in modo efficiente.

La base di gestione dei dati di Lenovo è costruita su un portafoglio di storage eterogeneo, tra cui:

- Soluzioni di storage tradizionali
- Storage definito da software (SDS)
- Offerte ad alte prestazioni da partner come Weka e DDN

Questa flessibilità si estende all'elaborazione dei dati, supportando sia i data lake open source che le soluzioni di livello aziendale come Cloudera. Per gli ambienti ibridi, i servizi professionali di Lenovo facilitano l'integrazione perfetta di offerte on-premise e basate sul cloud come Databricks.

© 2024 The Futurum Group. Tutti i diritti riservati.

Basandosi su questa solida base di dati, Lenovo ha fatto passi da gigante nello sviluppo di soluzioni di AI aziendale. Queste offerte sono progettate per essere di dimensioni adeguate, efficienti dal punto di vista energetico e affidabili, approvvigionandosi a varie applicazioni di AI, tra cui formazione, modellazione e inferenza.

Le caratteristiche principali dell'infrastruttura AI di Lenovo includono:

- Opzioni TruScale "as-a-service" (aaS) per: Gestione dati, Elaborazione ad alte prestazioni, GPU
- GPU as a Service (GPUaaS) che utilizza i server ThinkSystem (SR680a V3, SR685a e SR780a V3) che supportano fino a otto GPU ad alte prestazioni. Adatta per AI aziendale, calcolo ad alte prestazioni (HPC) e carichi di lavoro grafici

Questo approccio integrato garantisce che i clienti Lenovo possano gestire in modo efficiente i loro dati e distribuire senza problemi potenti soluzioni di AI, il tutto all'interno di soluzioni di infrastruttura flessibili e scalabili.

Lenovo integra l'AI nelle sue offerte per guidare i clienti nel loro percorso di adozione dell'AI. L'azienda ha ampliato la sua AI Services Practice con nuove soluzioni come AI Discover e AI Fast Start. Questi possono essere implementati come servizio tramite Lenovo TruScale, aiutando le aziende a distribuire e scalare le soluzioni di AI e GenAI in modo efficiente.

Per la gestione dell'infrastruttura, Lenovo offre XClarity One, una piattaforma cloud ibrida con un'interfaccia intuitiva per la gestione di server e dispositivi edge. Questa soluzione incorpora AI-Powered Smarter Support, che sfrutta l'apprendimento automatico per la manutenzione predittiva e l'automazione delle attività di routine. Inoltre, XClarity One presenta un hub di gestione sicura che utilizza l'AI generativa per adattare e migliorare continuamente le difese di sicurezza informatica.

Per rispondere alle preoccupazioni legate all'ambiente, Lenovo ha introdotto LISSA (Lenovo Intelligent Sustainability Solutions Advisor), uno strumento basato sull'AI che aiuta le aziende a ridurre l'impatto ambientale dell'IT. LISSA rappresenta un nuovo approccio alla gestione delle risorse IT durante il loro ciclo di vita, fornendo informazioni utili che aiutano le aziende a comprendere e gestire le loro impronte di carbonio in modo più efficiente.

Dal nostro punto di vista, la soluzione di raffreddamento a liquido Neptune® di sesta generazione di Lenovo è una scelta comprovata e leader di mercato per soddisfare i requisiti dei carichi di lavoro unici dell'AI. Questo metodo avanzato di raffreddamento a liquido supera notevolmente il tradizionale raffreddamento ad aria, in particolare nella dissipazione del calore, garantendo prestazioni migliori e migliorando sostanzialmente l'efficienza energetica. E i servizi di alimentazione e raffreddamento di Lenovo sono progettati per aiutare i data center a gestire il loro consumo energetico in modo più efficiente gestendo proattivamente i sistemi di alimentazione e raffreddamento per garantire operazioni affidabili e riducendo il rischio di tempi di inattività.

XClarity Energy Manager di Lenovo basato sull'AI aumenta l'ottimizzazione dell'efficienza energetica individuando i server poco utilizzati, spegnendo i componenti non utilizzat, e valutando come i server possono ospitare nuovi carichi di lavoro in base alle risorse disponibili. Inoltre, la piattaforma è ideale per sfruttare le analisi di guasto predittivo dell'AI sfruttando le funzionalità AI-Powered Smarter Support incluse in Lenovo XClarity One.

Lenovo Smarter AI for All sostiene la sua proposta di AI aziendale guidata da un investimento di 1 miliardo di dollari in tre anni, mirato ad accelerare l'implementazione dell'AI su più settori. Questo investimento supporta la creazione di oltre 150 soluzioni di AI in collaborazione con gli ISV. Lenovo offre una suite variegata di oltre 90 piattaforme di AI ibrida progettate per soddisfare una vasta gamma di esigenze nel settore pubblico, privato e personale.

Inoltre, le soluzioni di AI Lenovo hanno vinto numerosi premi, tra cui il CRN 2024 AI 100 e il CRN 2023 Product of the Year.

© 2024 The Futurum Group. Tutti i diritti riservati.

Fonte: Lenovo

La strategia ODM+ di Lenovo rafforza ulteriormente la sua offerta di AI aziendale. Questo approccio implica la proprietà da parte di Lenovo dei suoi processi di sviluppo e produzione, sfruttando la sua catena di fornitura da 20 miliardi di sollari per ridurre i costi e fornire servizi end-to-end. Il programma Lenovo AI Innovators è un ecosistema di partner software che lavorano con Lenovo per fornire ai clienti soluzioni di AI pronte per l'implementazione. Il programma si concentra sulla riduzione dei costi degli sviluppatori di software di AI e consente di accedere ai laboratori globali di AI per test e benchmarking. Il programma supporta gli ISV e i loro clienti nell'implementazione dell'AI su base scalabile e sicura.

© 2024 The Futurum Group. Tutti i diritti riservati.

Sezione 3: Conclusioni e raccomandazioni

Crediamo che la proposta Lenovo Smarter AI for All garantisca l'implementazione di soluzioni di AI aziendale pronte per il mercato. Lenovo offre il portafoglio e i servizi essenziali per guidare la collaborazione dei partner dell'ecosistema, le competenze in materia di AI, gli ISV, essenziali per il successo dell'AI aziendale a livello di organizzazione.

L'approccio di Lenovo garantisce che i clienti possano mantenere ed esercitare il controllo sui loro sistemi di AI, sull'infrastruttura di dati e sugli ambienti di distribuzione. L'approccio di Lenovo garantisce sicurezza e privacy dei dati, indipendentemente dalla posizione geografica. Con queste considerazioni, forniamo i seguenti consigli alle organizzazioni e agli ITBDM nella valutazione di Lenovo Smarter AI for All per soddisfare i loro requisiti di AI aziendale.

Funzionalità complete del portafoglio. Dai la priorità al fatto che Lenovo Smarter AI for All offre una vasta gamma di soluzioni di AI, dai dispositivi di AI personali all'infrastruttura AI di livello aziendale. Le soluzioni di AI aziendale Lenovo sono progettate per essere sicure, efficienti e scalabili, aiutando le aziende a utilizzare l'AI in modo responsabile e corretto.

Accelera i vantaggi dell'AI aziendale. I team Lenovo accelerano i processi attraverso il programma AI Innovators insieme agli ISV per fornire risultati aziendali mediante applicazioni AI costituite da soluzioni di gestione e protezione dei dati lavorando con partner, servizi e competenze di AI in grado di accelerare i processi di progettazione, strategia ed esecuzione. Le soluzioni AI Fast Start offrono casi d'uso e progetti pilota di settore con dati propri delle organizzazioni, nonché soluzioni di infrastruttura AI efficienti dal punto di vista energetico che offrono dimensioni adeguate e scelta.

Vantaggi di NVIDIA Alliance. Scopri come utilizzando la piattaforma software NVIDIA AI Enterprise, che include il framework NVIDIA NeMo, insieme al più recente design di riferimento GenAI di Lenovo basato su LLM, puoi distribuire e commercializzare strumenti e modelli di base GenAI sfruttando una soluzione pre-convalidata, completamente integrata e ottimizzata per le prestazioni. Grazie alla partnership con NVIDIA per collaborare allo sviluppo delle soluzioni, Lenovo può aiutare le aziende a migliorare i flussi di lavoro LLM e a massimizzare prestazioni, sicurezza e dimensionamento.

© 2024 The Futurum Group. Tutti i diritti riservati.

Informazioni importanti su questo report

COLLABORATORI

Steven Dickens
Chief Technology Advisor | The Futurum Group

Ron Westfall
Direttore della ricerca | The Futurum Group

EDITORE

Daniel Newman
CEO | The Futurum Group

DOMANDE

Contattaci se desideri discutere di questo report. The Futurum Group ti risponderà in breve tempo.

CITAZIONI

Questo articolo può essere citato da stampa e analisti accreditati, ma deve essere citato nel contesto, mostrando il nome dell'autore, il titolo dell'autore e "The Futurum Group". I soggetti che non sono addetti stampa o analisti hanno bisogno di previa autorizzazione scritta da parte di The Futurum Group per eventuali citazioni

LICENZE

Questo documento, inclusi tutti i materiali di supporto, è di proprietà di The Futurum Group. Questa pubblicazione non può essere riprodotta, distribuita o condivisa in alcuna forma senza la preventiva autorizzazione scritta di The Futurum Group.

INFORMATIVA

The Futurum Group fornisce ricerca, analisi e consulenza a molte aziende high-tech, comprese quelle menzionate in questo articolo. Nessun dipendente dell'azienda detiene posizioni azionarie presso le società citate in questo documento.

INFORMAZIONI SU LENOVO E NVIDIA

Lenovo offre a tutti la nuova era dell'innovazione basata sull'AI. Il nostro portafoglio full-stack offre soluzioni AI potenti, flessibili e responsabili per trasformare i settori ed emancipare le persone. Creiamo un futuro Smarter AI for all. Noi di Lenovo crediamo che il futuro dell'AI implichi la coesistenza di AI pubblica e aziendale. Lenovo offre AI a te e ai tuoi dati.

In collaborazione con NVIDIA, le soluzioni di AI ibrida sono appositamente costruite collaborando nella progettazione al fine di portare in modo efficiente l'AI ai dati dei clienti, dove e quando gli utenti ne hanno più bisogno, segnando un progresso nella visione di Lenovo di rendere possibile l'AI per tutti e offrendo supporto al time to market per tecnologie e architetture innovative per la prossima generazione di AI generativa. Le soluzioni ibride Lenovo, già ottimizzate per eseguire il software NVIDIA AI Enterprise per un'AI aziendale sicura, supportata e stabile, consentono inoltre agli sviluppatori di accedere ai microservizi NVIDIA, tra cui NIM NVIDIA e NeMo Retriever.

CHI SIAMO THE FUTURUM GROUP

The Futurum Group è una società indipendente di ricerca, analisi e consulenza, incentrata sull'innovazione digitale e sulle tecnologie e tendenze che rivoluzionano il mercato. Ogni giorno i nostri analisti, ricercatori e consulenti aiutano i leader aziendali di tutto il mondo ad anticipare i cambiamenti tettonici nei loro settori e a sfruttare l'innovazione dirompente per ottenere o mantenere un vantaggio competitivo nei loro mercati.

DATI DI CONTATTO

The Futurum Group LLC I futurumgroup.com I (833) 722-5337 I

© 2024 The Futurum Group. Tutti i diritti riservati.