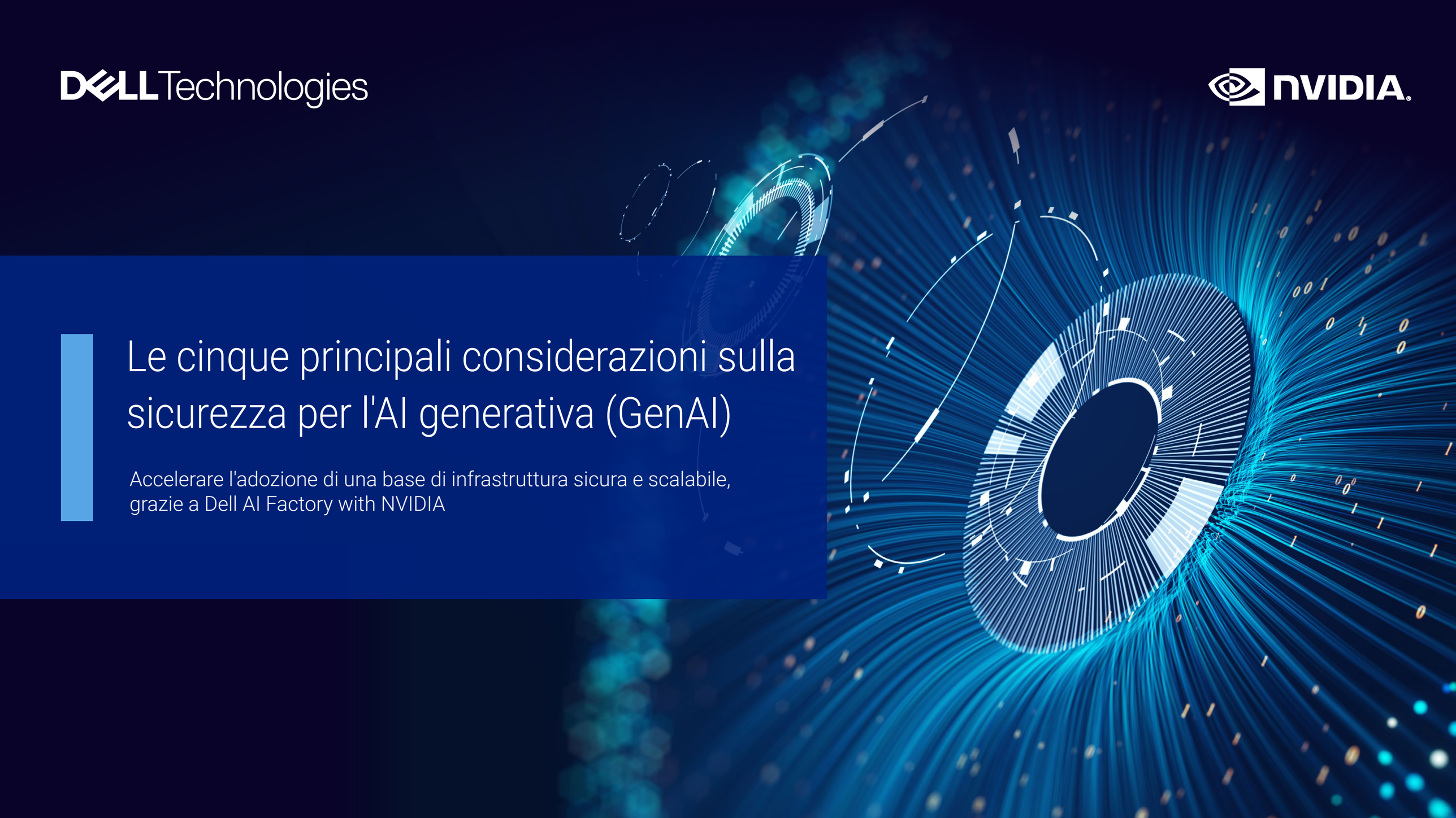
Le cinque principali considerazioni sulla sicurezza per l'AI generativa (GenAI)

Accelerare l'adozione di una base di infrastruttura sicura e scalabile, grazie a Dell AI Factory with NVIDIA



Il potenziale trasformativo della GenAI

La GenAI possiede potenzialità per cambiare il mondo che i visionari iniziano solo ora a immaginare.

Il 76%

dei responsabili IT e dei leader aziendali ritiene che la GenAI offrirà un valore trasformativo alla propria organizzazione[1]

AI

Analisi avanzata e tecniche basate sulla logica per interpretare gli eventi e supportare e automatizzare azioni e decisioni.

AI generativa

Tecnologie e tecniche che sfruttano grandi quantità di dati per generare nuovi contenuti da prompt in linguaggio naturale o da altri input non codificati e non tradizionali.

Simulazione

- Gemello digitale
- Dati sintetici
- Progettazione di framework
- Previsione

Creazione di contenuti

- Codifica
- Matematica
- Scrittura e discorso
- Immagini e video
- Audio

Scoperta di contenuti

- Ricerca con linguaggio naturale
- Analisi di dataset di grandi dimensioni
- Gestione delle conoscenze
- Istruzione e formazione personalizzate

Esperienza utente

- Traduzioni in tempo reale per oltre 70 lingue
- Interazioni personalizzate utilizzando espressioni facciali naturali e linguaggio del corpo

1 Studio Dell Technologies Innovation Catalyst, febbraio 2024

Maggiore potenziale, maggiore rischio

È allettante per i leader aziendali potersi muoversi rapidamente, ignorando le implicazioni relative a dati, conformità, governance e altri rischi. Ma la GenAI è una spada a doppio taglio quando si tratta di sicurezza.

Vantaggi

- Rilevamento delle minacce migliorato
- Maggiore efficienza operativa
- Formazione personalizzata per la sensibilizzazione alla sicurezza

Svantaggi

- Attacchi più sofisticati
- Social engineering avanzato
- Shadow AI

Il 33%

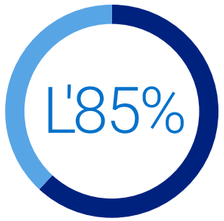
degli intervistati ha indicato la sicurezza informatica come il principale rischio della GenAI che le loro organizzazioni stanno cercando di mitigare.[2]

2 Indagine globale McKinsey sull'IA: The state of AI in early, maggio 2024

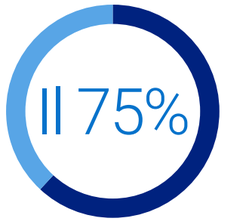
CONSIDERAZIONE 1

Il nuovo panorama delle minacce

Insieme alla promessa della GenAI arriva una realtà inquietante: gli autori di attacchi stanno creando minacce nuove e più complesse, in grado di aggirare le difese convenzionali, rendendo difficile per i team di sicurezza informatica stare al passo.



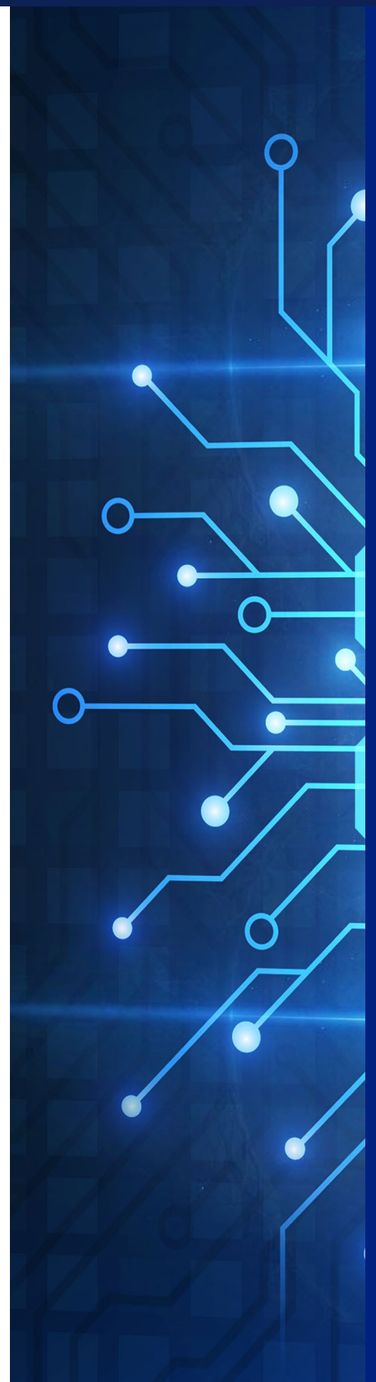
degli intervistati ritiene che l'AI abbia reso più sofisticati gli attacchi alla sicurezza informatica.[3]



dei professionisti della sicurezza ha registrato un aumento degli attacchi negli ultimi 12 mesi.[4]

Per proteggersi da queste minacce emergenti, le aziende devono concentrarsi sulla riduzione al minimo della superficie di attacco, ad esempio attraverso test di penetrazione, monitoraggio e verifica.

3 2024 Human Risk in Cybersecurity Survey, EY, maggio 2024
4 Voice of SecOps Report "Generative AI and Cybersecurity: Bright Future o Business Battleground?" 2023



Vettori di attacco emergenti

Malware avanzato
Malware sempre più sofisticato che utilizza l'AI generativa per "evolversi autonomamente", cambiando continuamente il codice per non essere rilevato dalla sicurezza esistente, come il rilevamento basato su firma.

E-mail e campagne di phishing altamente personalizzate
Si registra un aumento della frequenza delle e-mail malevoli con aspetto autentico, che non presentano i consueti segnali di truffa.

Dati deep fake convincenti Furti di identità, frodi finanziarie e disinformazione resi più semplici dalla capacità di imitare azioni umane, come scrittura, discorso, immagini o video.

Ricognizione automatizzata
Raccolta di informazioni per identificare vulnerabilità e punti deboli nella rete o nel sistema di un potenziale obiettivo per facilitare attacchi più mirati.

CONSIDERAZIONE 2

Rischi di deployment e implementazione

Le organizzazioni che vogliono sfruttare i potenziali benefici della GenAI hanno bisogno di grandi quantità di dati di alta qualità – input che i modelli possono utilizzare per produrre i migliori risultati. Ma dati e rischio vanno di pari passo. Prima di sfruttare qualsiasi informazione, le aziende devono valutare e prendere in considerazione i loro requisiti unici, gli input e i rischi.

Vulnerabilità dei modelli linguistici di grandi dimensioni (LLM)

I servizi di GenAI sono vulnerabili agli attacchi di injection rapida, in cui gli autori di attacchi manipolano gli output per aggirare i guardrail di sicurezza od ottenere accesso non autorizzato ai file che potrebbero essere stati utilizzati per perfezionare il modello.

Avvelenamento dei dati

Gli autori di attacchi possono inviare deliberatamente dati alterati a un LLM durante la fase di addestramento. Ciò può rendere il modello vulnerabile agli attacchi tramite backdoor integrati nei dati. Un esempio reale è l'attacco e lo sfruttamento dei filtri di spam, addestrandoli sulle e-mail di spam.

Complessità normativa

Le autorità di regolamentazione di tutto il mondo si battono per comprendere, controllare e garantire la sicurezza della GenAI. Sebbene i modelli di GenAI siano soggetti alle attuali regole di sovranità dei dati, che determinano il modo in cui i dati vengono archiviati, elaborati e utilizzati, gli enti governativi stanno ancora definendo la supervisione della proprietà intellettuale e delle informazioni protette da copyright. Il rispetto delle normative può risultare oneroso, ma il mancato adempimento alle normative già esistenti e di quelle in corso di elaborazione potrebbe comportare multe e altre sanzioni.

CONSIDERAZIONE 3

Shadow AI

Molti dipendenti oggi utilizzano già generatori pubblici di testo, immagini e video, come ChatGPT, per aumentare i flussi di lavoro quotidiani. Tuttavia, quando questi strumenti vengono utilizzati senza una governance adeguata, rappresentano una minaccia critica per le organizzazioni che cercano di proteggere la proprietà intellettuale e i dati aziendali. Questo utilizzo non autorizzato della GenAI è noto come Shadow AI.

Perdita di proprietà intellettuale

Le aziende si stanno già occupando della perdita di proprietà intellettuale dovuta alla condivisione di informazioni sensibili da parte dei dipendenti negli strumenti pubblici di GenAI.

Perdita di dati del codice sorgente

Gli sviluppatori che tentano di ottimizzare il codice sorgente utilizzando ChatGPT hanno causato perdite di dati.

Per affrontare le sfide della Shadow AI, le società dovrebbero implementare un comitato o un consiglio a livello aziendale, conferendogli l'autorità di prendere decisioni che prevedono una governance sicura dell'AI.

Dove risiedono i dati? Dove devono essere posizionati i carichi di lavoro?

L'AI raggiunge il massimo delle sue potenzialità quando è abbinata ai dati, ovunque essi risiedano. Con il controllo completo sull'infrastruttura e sugli LLM, non vi è alcun rischio di perdita di IP o di dati del codice sorgente.

Costi

L'utilizzo ottimale delle implementazioni on-premise può ridurre il TCO fino al 75% in 3 anni.[5]

Sicurezza e privacy

Creare ambienti di AI/GenAI sicuri in tutta l'organizzazione, con flussi di lavoro e operazioni on-premise. Esercitare un rigoroso controllo sulla sicurezza dei dati e sul rispetto delle normative di conformità, in particolare per i settori che gestiscono dati sensibili.

5 Dati basati su una ricerca di Enterprise Strategy Group commissionata da Dell, che confronta l'infrastruttura Dell on-premise con l'Infrastructure as-a-Service su public cloud nativa, aprile 2024. I modelli analizzati mostrano che un LLM con parametro 7B che utilizza RAG è fino al 38% più conveniente per un'organizzazione di 5.000 utenti, mentre un LLM con parametro 70B che utilizza RAG è fino al 75% più conveniente per un'organizzazione di 50.000 utenti. I risultati effettivi possono variare. Sintesi economica

CONSIDERAZIONE 4

Criteri di valutazione

Nell'ultimo anno, la community AI si è sempre più concentrata su tre questioni chiave: sviluppo e deployment responsabili, valutazione dell'impatto e riduzione dei rischi. Quando le aziende esaminano i modelli di GenAI, devono prestare attenzione ad alcune avvertenze fondamentali:

Assenza di requisiti di reporting coerenti

Gli sviluppatori leader testano principalmente i propri modelli rispetto a diversi benchmark AI responsabili. A causa della significativa mancanza di standardizzazione nella presentazione dei dati, risulta difficile confrontare metodicamente i rischi e le limitazioni dei principali modelli di AI.

Materiale protetto da copyright negli output

Gli output degli LLM più diffusi possono contenere materiale protetto da copyright, violando potenzialmente la legge ed esponendo le aziende che utilizzano tale materiale al rischio di sanzioni.

Le vulnerabilità sono sempre più complesse

I ricercatori stanno trovando strategie meno ovvie che causano un comportamento dannoso degli LLM, come chiedere ai modelli di ripetere all'infinito parole casuali.

Gli sviluppatori mancano di trasparenza

In molti casi, gli sviluppatori AI non sono cooperativi per quanto riguarda i loro dati e le loro metodologie di addestramento. Questo rende più difficile comprendere in modo più approfondito la robustezza e la sicurezza dei sistemi di intelligenza artificiale.

CONSIDERAZIONE 5

Vantaggi per la sicurezza

Accanto ai rischi per la sicurezza della GenAI, vi sono anche potenziali vantaggi. Schiudendo nuove possibilità di protezione, la GenAI sta diventando un'alleata fondamentale in fatto di sicurezza informatica.

Ora è possibile iniziare a creare operazioni di sicurezza scalabili con accesso più rapido a informazioni approfondite e rilevamento automatico delle minacce, offrendo efficienza e integrando team di sicurezza sottodimensionati in termini di personale.

Threat Detection and Response

Analizzando i dati storici e identificando modelli e anomalie, la GenAI è in grado di riconoscere minacce nuove e in continua evoluzione in tempo reale. Può monitorare costantemente il traffico di rete, i log del sistema e il comportamento degli utenti, identificando immediatamente attività irregolari che potrebbero segnalare minacce alla sicurezza.

Il risultato è un potente rilevamento adattivo delle minacce, che consente una rapida risposta ai mutevoli vettori di attacco e fornisce un meccanismo di difesa proattivo contro le minacce informatiche emergenti.

Simulazioni e formazione per la gestione delle minacce

Grazie all'intelligenza artificiale generativa, le aziende sono in grado di simulare una vasta gamma di minacce alla sicurezza informatica e scenari di attacco in un ambiente controllato. Di conseguenza, i team sono meglio preparati a identificare, rispondere e mitigare le minacce informatiche quando la reattività è essenziale.

Analisi e riepiloghi approfonditi

La GenAI consente ai team di analizzare i dati provenienti da fonti o moduli diversi, permettendo loro di eseguire analisi dei dati tradizionalmente lunghe e noiose in modo più rapido e accurato. I team possono inoltre creare riepiloghi in linguaggio naturale degli incidenti e delle valutazioni delle minacce, migliorando l'efficienza e aumentando il risultato del team.

Formazione personalizzata per la sensibilizzazione alla sicurezza

Integrando l'AI conversazionale in aggiunta alla GenAI e inserendo un avatar AI nell'interfaccia utente, le organizzazioni possono fornire interazioni personalizzate (disponibili su scala 24/7), utilizzando espressioni facciali e linguaggio del corpo naturali. Ciò può essere utilizzato per la formazione e l'istruzione sulla sicurezza, offrendo un'esperienza di apprendimento più naturale, personalizzata e interattiva, valutazioni automatizzate e molto altro ancora.

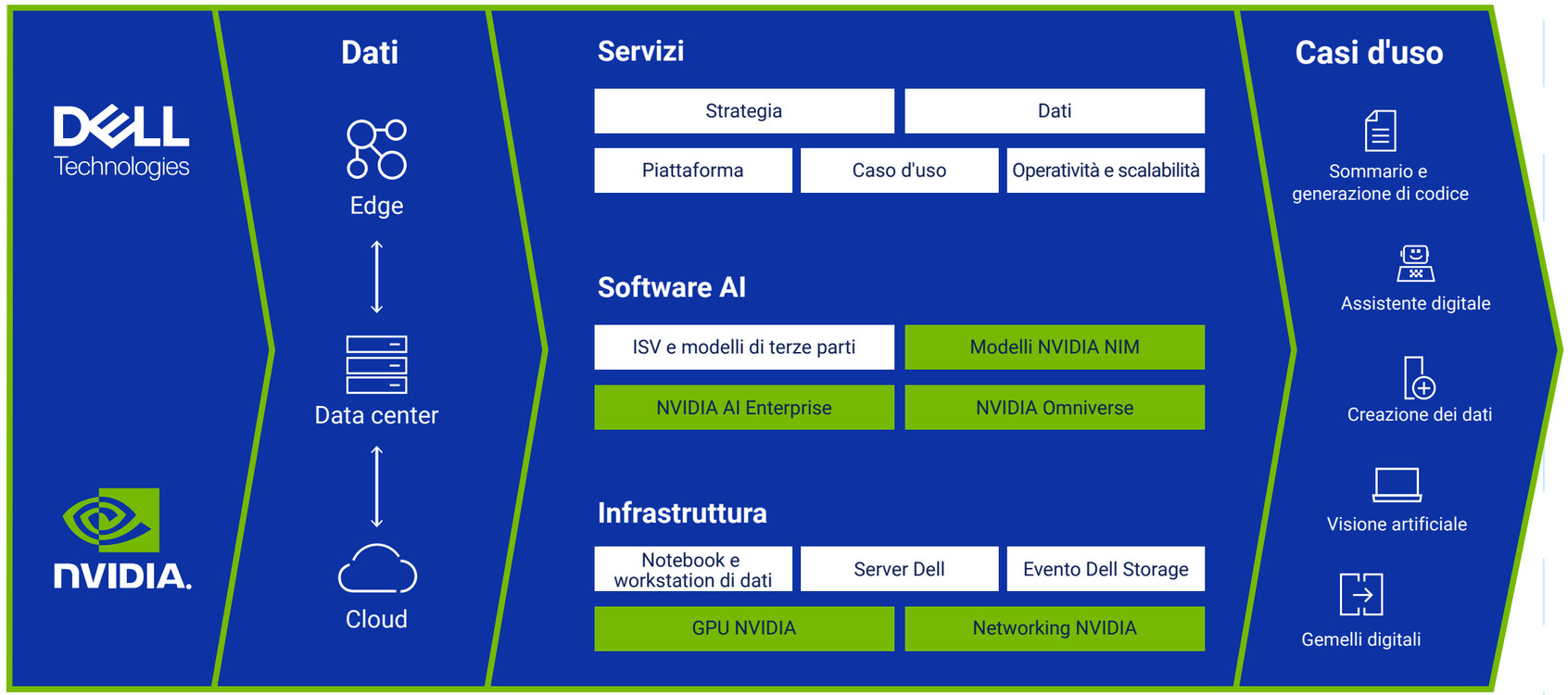
Dell AI Factory with NVIDIA

Accelera il percorso verso l'AI e trasforma in modo sicuro i dati in informazioni approfondite con la prima soluzione di AI completa e pronta all'uso del settore. Dell AI Factory with NVIDIA risponde alle complesse esigenze delle aziende che desiderano sfruttare l'AI e la GenAI. Con un'infrastruttura e servizi leader del settore, insieme al software AI di NVIDIA, è possibile aumentare il time-to-value dei progetti, semplificando lo sviluppo e il deployment.

- Ridurre il rischio di compromissione con un'infrastruttura dotata di sicurezza intrinseca, tra cui root-of-trust e altre funzionalità chiave.
- Proteggere i dati da fughe che potrebbero causare la perdita di proprietà intellettuale con una soluzione di AI on-premise controllata.
- Soddisfare i rigorosi requisiti di conformità e sovranità dei dati, introducendo l'AI nei dati con accesso sicuro.
- Proteggere la privacy delle entità interessate, controllando dove e chi ha accesso ai dati.

Dell AI Factory with NVIDIA

LA PRIMA SOLUZIONE DI AI END-TO-END PER AZIENDE DEL SETTORE



I dati alimentano AI Factory e i casi d'uso dell'utente

I dati più preziosi si trovano on-premise e nell'edge. Dell Technologies contribuisce all'integrazione dell'AI nei preziosi dati aziendali ed è leader nello storage, nella protezione e nella gestione di tali dati.

Dai casi d'uso ai risultati

L'AI Factory produce risultati di business basati sui casi d'uso prioritari. Dell Technologies semplifica il deployment dei casi d'uso AI più importanti con soluzioni convalidate e servizi su misura.

Non lasciate che i rischi per la sicurezza ostacolino l'innovazione

Ti aiutiamo a esplorare il mondo dell'AI e della GenAI, affinché tu possa coglierne tutti i vantaggi.

PIANIFICAZIONE STRATEGICA

Accelerator Workshop for GenAI gratuito

- Per iniziare il percorso verso lo sviluppo di una strategia vincente
- Affrontare sfide e lacune, definire gli obiettivi prioritari e identificare le opportunità
- Ottenere una valutazione dell'idoneità per un'analisi più approfondita dei requisiti dell'infrastruttura, dei modelli di AI, delle integrazioni operative e molto altro ancora

PREPARAZIONE TECNICA

Un laboratorio mobile pronto all'uso

Inizia il tuo percorso verso il successo. Include un Dell Mobile Precision Workstation 5690 / 7780 con GPU NVIDIA e due giorni di servizi di consulenza per aiutarti a iniziare.

- Ambiente sandbox portatile per test e dimostrazioni di GenAI
- Pre-convalidato con la piattaforma NVIDIA AI Workbench, già pronta per gli sviluppatori
- Primo caso d'uso del chatbot implementato utilizzando i tuoi dati
- Un approccio conveniente e a basso rischio per sperimentare e sviluppare le competenze di GenAI

DELL MOBILE PRECISION WORKSTATION 5690 / 7780 CON GPU NVIDIA

INIZIA SUBITO

DELL Technologies AI Factory WITH NVIDIA