DELL Technologies

NVIDIA

5 principales considérations de sécurité pour l'IA générative

Accélérez votre adoption d'une infrastructure sécurisée et évolutive avec Dell AI Factory with NVIDIA

Potentiel de transformation de l'IA générative

L'IA générative a le potentiel de changer la donne d'une manière que les visionnaires commencent à peine à imaginer.

76 %

des responsables IT et métier pensent que l'IA générative apportera une valeur transformationnelle à leur organisation.[1]

IA

Analyse avancée et techniques logiques qui interprètent les événements, mais aussi soutiennent et automatisent les décisions et les actions.

GenAI

Technologies et techniques qui utilisent des quantités importantes de données pour créer du nouveau contenu à partir d'instructions en langue naturelle ou d'autres entrées non codées et non traditionnelles.

Simulation

- Jumeau numérique
- Données synthétiques
- Cadres de conception
- Prédiction

Création de contenu

- Codage
- Mathématiques
- Voix/écriture
- Images et vidéos
- Audio

Découverte de contenu

- Recherche en langage naturel
- Analyse d'ensembles de données volumineux
- Gestion des connaissances
- Apprentissage et formation personnalisés

Expérience utilisateur

- Traductions en temps réel pour plus de 70 langues
- Interactions personnalisées utilisant des expressions faciales naturelles et un langage corporel

1 Dell Technologies, étude « Innovation Catalysts », février 2024

Potentiel accru, risque accru

Il est tentant pour les dirigeants d'entreprise de vouloir agir rapidement, en contournant les implications liées aux données, à la conformité, à la gouvernance et à d'autres risques. L'IA générative est une arme à double tranchant en matière de sécurité.

Avantages

- Détection améliorée des menaces
- Amélioration de l'efficacité opérationnelle
- Formation personnalisée de sensibilisation à la sécurité

Inconvénients

- Sophistication renforcée des attaques
- Ingénierie sociale avancée
- IA fantôme

33 %

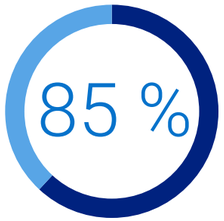
des personnes interrogées ont indiqué que la cybersécurité était le principal risque lié à l'IA générative que leur organisation s'efforce d'atténuer.[2]

2 McKinsey Global Survey on AI: The state of AI in early, May 2024

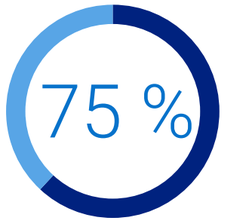
CONSIDÉRATION Nº 1

Nouveau paysage des menaces

À la promesse de l'IA générative s'ajoute une réalité qui fait réfléchir : les pirates créent de nouvelles attaques plus complexes capables de contourner les défenses conventionnelles, ce qui peut empêcher les équipes de cybersécurité de suivre le rythme.



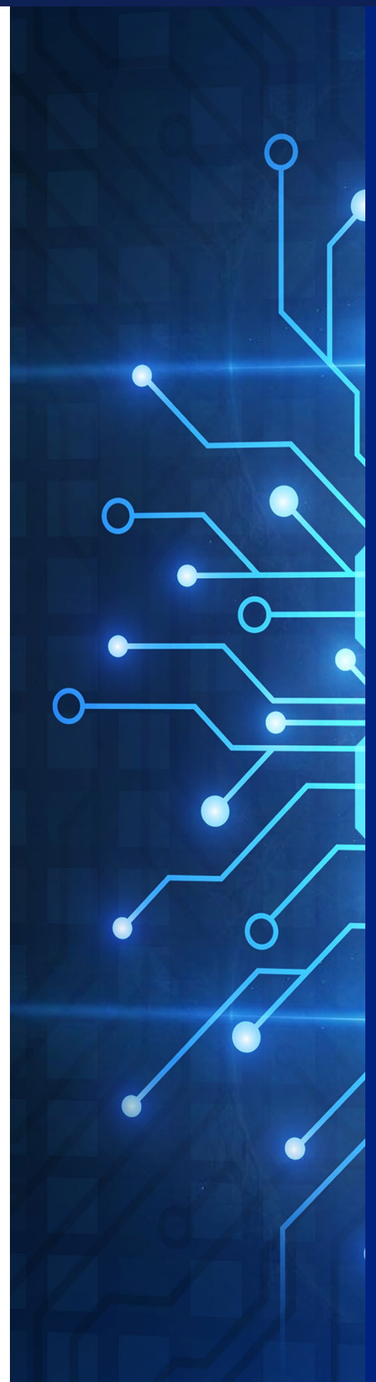
des personnes interrogées pensent que l'IA a rendu les attaques de cybersécurité plus sophistiquées.[3]



des professionnels de la sécurité ont constaté une augmentation des attaques au cours des 12 derniers mois.[4]

Pour se protéger contre ces menaces émergentes, les entreprises doivent se concentrer sur la réduction de la surface d'attaque en effectuant des tests d'intrusion, une surveillance et des audits, par exemple.

3 2024 Human Risk in Cybersecurity Survey, EY, May 2024
4 Rapport Voice of SecOps « Generative AI and Cybersecurity: Bright Future or Business Battleground? » 2023



Vecteurs d'attaque émergents

Logiciel malveillant avancé
Logiciel malveillant de plus en plus sophistiqué qui utilise l'IA générative pour « auto-évoluer », modifiant continuellement son code pour qu'il ne soit pas détecté par la sécurité existante (p. ex. la détection basée sur les signatures).

E-mails et campagnes de phishing hautement personnalisés
Fréquence croissante d'e-mails malveillants d'apparence authentique qui ne présentent pas de signes d'arnaque habituels.

Données deepfake convaincantes
Vol d'identité, fraude financière et désinformation facilités par la possibilité d'imiter des actions humaines, telles que l'écriture, la parole, les images ou la vidéo.

Reconnaissance automatisée
Collecte d'informations qui identifie les failles et les vulnérabilités du réseau ou du système d'une cible potentielle afin de faciliter des attaques plus ciblées.

CONSIDÉRATION Nº 2

Risques liés au déploiement et à l'implémentation

Les entreprises qui souhaitent mettre à profit les avantages potentiels de l'IA générative ont besoin de grandes quantités de données de haute qualité, des entrées que les modèles peuvent utiliser pour produire les meilleurs résultats. Les données et les risques vont de pair. Avant d'exploiter des informations, les entreprises doivent évaluer soigneusement leurs exigences, leurs entrées et leurs risques uniques et en tenir compte.

Vulnérabilités des grands modèles de langage (LLM)

Les services d'IA générative sont vulnérables aux attaques par injection rapide, au cours desquelles les pirates manipulent les résultats pour contourner les barrières de sécurité ou obtenir un accès non autorisé aux fichiers ayant pu être utilisés pour affiner le modèle.

Empoisonnement des données

Les pirates peuvent délibérément fournir des données modifiées à un LLM lors de la phase de formation. Cela peut rendre le modèle vulnérable aux attaques par le biais de portes dérobées intégrées dans les données. Exemple concret : l'attaque et l'exploitation des filtres anti-spam en les entraînant aux e-mails indésirables.

Complexité réglementaire

Les organismes de réglementation du monde entier se battent pour comprendre, contrôler et garantir la sécurité de l'IA générative. Même si les modèles d'IA générative sont soumis aux règles actuelles de souveraineté des données qui dictent la façon dont les données sont stockées, traitées et utilisées, les organes dirigeants continuent de définir la surveillance des propriétés intellectuelles et des informations protégées par le droit d'auteur. Respecter les réglementations peut s'avérer coûteux, mais le non-respect des réglementations établies et émergentes peut entraîner des amendes et d'autres pénalités.

CONSIDÉRATION Nº 3

IA fantôme

De nombreux collaborateurs utilisent déjà des générateurs de texte, d'images et de vidéos publics tels que ChatGPT pour augmenter leurs flux de travail quotidiens. Toutefois, lorsque ces outils sont utilisés sans gouvernance appropriée, ils constituent une menace critique pour les entreprises qui tentent de sécuriser la propriété intellectuelle et les données d'entreprise. Cette utilisation non autorisée de l'IA générative est connue sous le nom d'IA fantôme.

Perte de propriété intellectuelle

Dès à présent, les entreprises font face à la perte de propriété intellectuelle du fait des employés qui partagent des données sensibles en public dans des outils d'IA.

Fuite de données de code source

Les développeurs tentant d'optimiser le code source à l'aide de ChatGPT ont provoqué des fuites de données.

Pour relever les défis de l'IA fantôme, les entreprises doivent mettre en place un conseil ou un bureau à l'échelle de l'entreprise habilité à prendre des décisions impliquant une gouvernance sécurisée de l'IA.

Où se trouvent vos données ? Où placer les charges applicatives ?

L'IA fonctionne mieux lorsqu'elle est associée à vos données, où qu'elles se trouvent. Avec un contrôle total sur l'infrastructure et les LLM, il n'y a aucun risque de perte d'adresse IP ou de fuite de données de code source.

Coûts

L'utilisation d'implémentations sur site peut réduire le coût total de possession jusqu'à 75 % sur 3 ans.[5]

Sécurité et confidentialité

Sécurisez les environnements d'IA/IA générative dans l'ensemble de l'entreprise avec des workflows et des opérations sur site. Exercez un contrôle strict sur la sécurité des données et le respect des réglementations de conformité, en particulier pour les secteurs qui gèrent des données sensibles.

5. D'après une étude Enterprise Strategy Group réalisée à la demande de Dell, comparant l'infrastructure Dell sur site à l'infrastructure as-a-service de Cloud public native, avril 2024. Les modèles analysés montrent qu'un LLM à 7 milliards de paramètres utilisant la RAG pour une organisation de 5 000 utilisateurs est jusqu'à 38 % plus rentable tandis qu'un LLM à 70 milliards de paramètres utilisant la RAG pour une organisation de 50 000 utilisateurs est jusqu'à 75 % plus rentable. Les résultats réels peuvent varier. Synthèse économique

CONSIDÉRATION Nº 4

Critères d'évaluation

Au cours de l'année écoulée, la communauté de l'IA s'est de plus en plus concentrée sur trois questions clés : le développement et le déploiement responsables, l'évaluation de l'impact et l'atténuation des risques. À mesure que les entreprises évaluent les modèles d'IA générative, elles doivent prendre en compte certaines mises en garde importantes :

Pas d'exigences cohérentes en matière de reporting

Les principaux développeurs testent essentiellement leurs modèles par rapport à différentes références en matière d'IA responsable. En raison de ce manque significatif de standardisation des rapports, il est difficile de comparer méthodiquement les risques et les limites des principaux modèles d'IA.

Matériel protégé par des droits d'auteur dans les sorties

Les résultats des LLM populaires peuvent contenir du matériel protégé par des droits d'auteur, ce qui peut potentiellement enfreindre la loi et exposer les entreprises qui utilisent le matériel à des sanctions.

Les failles de sécurité sont de plus en plus complexes

Les chercheurs trouvent des stratégies moins évidentes qui provoquent des comportements de LLM nuisibles, comme demander aux modèles de répéter infiniment des mots aléatoires.

Les développeurs manquent de transparence

Dans de nombreux cas, les développeurs d'IA ne sont pas ouverts en ce qui concerne leurs données et méthodologies d'entraînement. Cela entrave les efforts visant à mieux comprendre la robustesse et la sécurité des systèmes d'IA.

CONSIDÉRATION Nº 5

Avantages en matière de sécurité

À côté des risques de sécurité de l'IA générative, il y a ses avantages potentiels en matière de sécurité. L'IA générative est devenue un allié essentiel dans le domaine de la cybersécurité, car elle offre de nouvelles opportunités en matière de protection.

Vous pouvez désormais commencer à créer des opérations de sécurité évolutives fournissant un accès plus rapide à des informations plus riches et une détection automatique des menaces, ce qui offre une efficacité accrue et constitue un soutien précieux pour les équipes de sécurité en sous-effectif.

Threat Detection and Response

En analysant les données historiques et en identifiant les schémas et les anomalies, l'IA générative peut reconnaître les menaces nouvelles et en constante évolution en temps réel. Elle peut surveiller en continu le trafic réseau, les journaux système et le comportement des utilisateurs, et identifier rapidement les activités irrégulières susceptibles de représenter des menaces de sécurité.

Le résultat est une détection des menaces très adaptative, qui permet une réponse rapide à l'évolution des vecteurs d'attaque et fournit un mécanisme de défense proactif contre les cybermenaces émergentes.

Formation et simulation de menaces

Avec l'IA générative, les entreprises peuvent simuler un large éventail de menaces de cybersécurité et de scénarios d'attaque dans un environnement contrôlé. Par conséquent, les équipes sont mieux préparées à identifier les cybermenaces, à y répondre et à les atténuer lorsque le temps compte.

Analyse et résumé approfondis

L'IA générative permet aux équipes d'enquêter sur les données provenant de différentes sources ou modules, ce qui leur permet d'effectuer des analyses de données traditionnellement chronophages et fastidieuses de manière plus rapide et précise. Les équipes peuvent également créer des résumés en langage naturel des incidents et des évaluations des menaces, ce qui améliore l'efficacité et augmente le rendement de l'équipe.

Formation personnalisée de sensibilisation à la sécurité

En intégrant l'IA conversationnelle à l'IA générative et en incorporant un avatar IA dans l'interface utilisateur, les organisations peuvent proposer des interactions personnalisées (disponibles à grande échelle 24 h/24 et 7 j/7) à l'aide d'expressions faciales naturelles et d'un langage corporel. Elle peut être utilisée pour la formation et l'éducation à la sécurité, offrant une expérience d'apprentissage plus naturelle, personnalisée et interactive, des évaluations automatisées, etc.

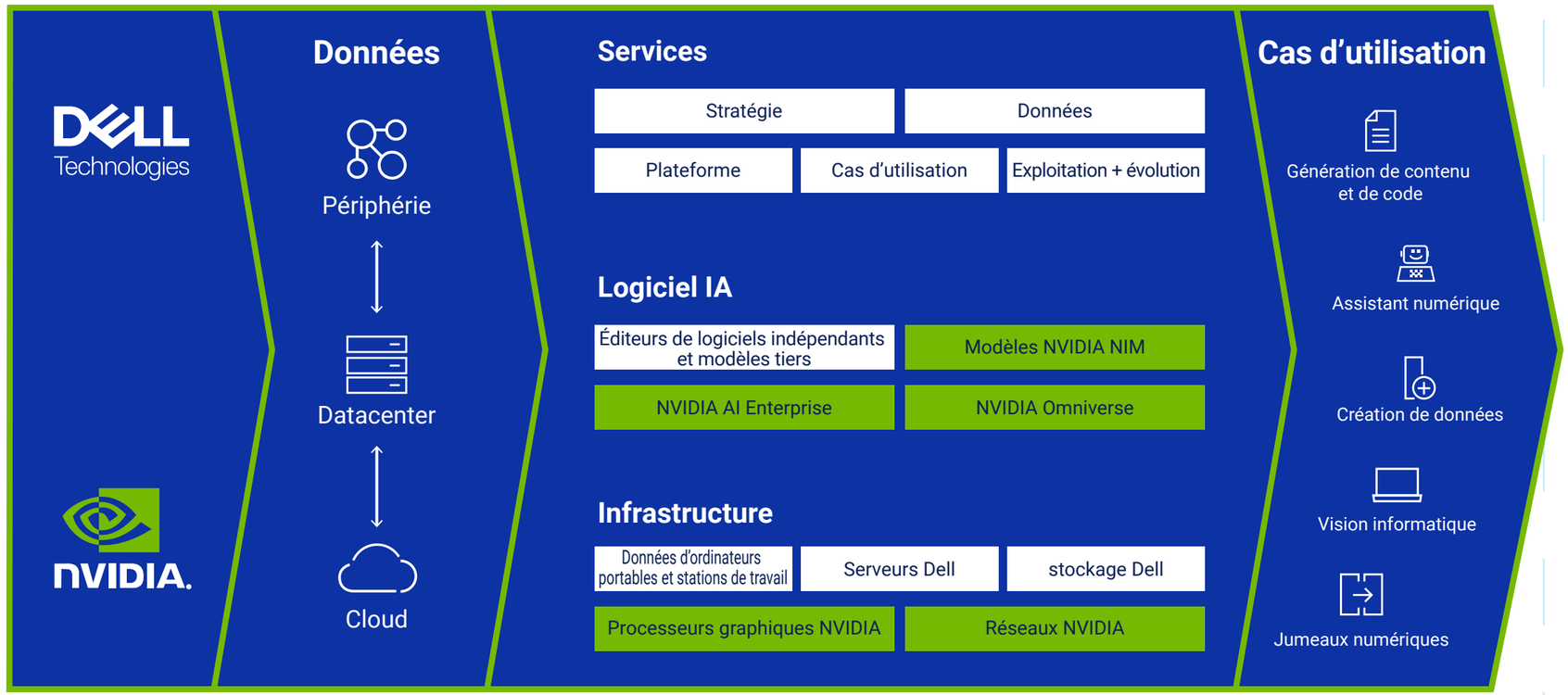
Dell AI Factory with NVIDIA

Accélérez votre transition vers l'IA et transformez en toute sécurité vos données en informations exploitables grâce à la première solution d'IA complète et clé en main du secteur. Dell AI Factory with NVIDIA répond aux besoins complexes des entreprises qui cherchent à tirer parti de l'IA et de l'IA générative. Grâce à une infrastructure et des services de pointe, associés au logiciel NVIDIA AI, vous pouvez accélérer le délai de rentabilisation de vos projets en simplifiant le développement et le déploiement.

- Réduisez le risque de compromis en utilisant une infrastructure qui intègre des mesures de sécurité intrinsèques, telles que la racine de confiance et d'autres caractéristiques essentielles.
- Protégez vos données contre les fuites qui pourraient entraîner une perte de propriété intellectuelle avec une solution d'IA sur site que vous contrôlez.
- Respectez les exigences strictes en matière de conformité et de souveraineté des données en apportant l'IA à vos données avec un accès sécurisé.
- Protégez la vie privée de vos parties prenantes en contrôlant où et qui peut accéder à vos données.

Dell AI Factory with NVIDIA

LA PREMIÈRE SOLUTION D'IA D'ENTREPRISE DE BOUT EN BOUT DU SECTEUR



Les données alimentent votre usine IA et vos cas d'utilisation

Vos données les plus précieuses se trouvent sur site et en périphérie. Dell Technologies vous aide à utiliser l'IA pour ces données précieuses. De plus, Dell est un leader en matière de stockage, de protection et de gestion de ces mêmes données.

Cas d'utilisation des résultats

L'AI Factory génère des résultats opérationnels optimisés par vos cas d'utilisation prioritaires. Dell Technologies simplifie le déploiement de vos cas d'utilisation de l'IA les plus importants avec des solutions validées et des services sur mesure.

Ne laissez pas les risques de sécurité entraver l'innovation

Nous vous guidons à travers le vaste univers de l'IA et de l'IA générative afin que vous puissiez profiter pleinement des avantages qu'elles offrent.

PLANIFICATION STRATÉGIQUE

Accelerator Workshop for GenAI (gratuit)

- Développez une stratégie gagnante
- Relever les défis, combler les lacunes, hiérarchiser les objectifs et identifier les opportunités
- Vous pouvez obtenir une évaluation de l'état de préparation pour mieux comprendre les exigences en matière d'infrastructure, les modèles d'IA, les intégrations opérationnelles, etc.

PRÉPARATION TECHNIQUE

Laboratoire mobile prêt à l'emploi

Lancez-vous sur la voie de la réussite. Comprend une station de travail mobile Dell Precision 5690/7780 avec des processeurs graphiques NVIDIA et deux jours de services de consultation pour vous aider à démarrer.

- Environnement sandbox portable pour les tests et la démonstration de l'IA générative
- Prévalidé avec la plateforme NVIDIA AI Workbench prête pour les développeurs
- Cas d'utilisation du chatbot initial implémenté avec vos données
- Approche à faible risque et à coût réduit pour tester et développer les compétences de l'IA générative

STATION DE TRAVAIL MOBILE DELL PRECISION 5690/7780 AVEC PROCESSEURS GRAPHIQUES NVIDIA

LANCEZ-VOUS DÈS MAINTENANT

DELL Technologies AI Factory WITH NVIDIA