DELL Technologies

5 Recommandations pour maximiser la valeur de l'IA générative en toute sécurité

1	2	3	4	5
Sécuriser les couches d'un système d'IA générative	**Utiliser les principes Zero-Trust**	**Assurer la gouvernance et la supervision humaine**	**Profiter des outils de sécurité basés sur l'IA générative dès qu'ils sont disponibles**	**Innover en toute confiance**
Infrastructure	Confiance zéro, vérification systématique	Impliquer les principales parties prenantes	Contenu	Une cybersécurité au service de l'entreprise
Système d'exploitation et Kubernetes	Accès du moindre privilège	Définir des règles de conformité éthique et réglementaire, et de gestion des données	Prédiction des risques	Redonner à l'entreprise foi en l'innovation grâce à une cybersécurité mature
Applications d'IA générative	Renforcement du système	Surveiller et assurer la responsabilité de chacun	Connaissances et automatisation	
Données	Gestion des identités	Formation et sensibilisation		
	Segmentation			
	Journalisation, surveillance et audit			

Si elle offre un réel potentiel de transformation à l'entreprise, la technologie d'IA générative présente des défis de sécurité inédits.

L'IA générative révolutionne l'entreprise comme jamais auparavant, en favorisant l'innovation et en offrant des avantages inégalés qui confèrent un réel avantage concurrentiel. Et même si cette technologie a le potentiel pour transformer l'entreprise, elle présente également des défis inédits en matière de sécurité.

Steve Brodson, Services Product Manager et Eitan Lederman, Cybersecurity Consultant, tous deux experts en la matière chez Dell, ont rejoint Chris Cicotte de l'équipe marketing APEX et IA pour répondre à ces préoccupations et discuter des moyens d'optimiser l'IA générative en toute sécurité. Lisez la suite pour obtenir un résumé de la conversation et des informations supplémentaires sur le sujet. Vous retrouverez l'intégralité de l'échange sur dell.com/cybersecuritymonth.

> « Les utilisateurs doivent être formés. Ils doivent savoir comment utiliser le système d'IA générative, ce qu'il faut faire, mais aussi ce qu'il ne faut pas faire. »
>
> ***Eitan Lederman***
> *Consultant en cybersécurité Dell*

Sécuriser les couches d'un système d'IA générative

Si l'IA générative est une technologie relativement nouvelle, la plupart des protocoles de sécurité correspondent aux techniques de cybersécurité utilisées pour sécuriser d'autres charges applicatives.

Infrastructure : concentrez-vous sur la réduction de la surface d'exposition aux attaques :

- Tests de vulnérabilité et de pénétration
- Application de correctifs
- Renforcement
- Gestion des identités, y compris les mots de passe forts, authentification multifacteur
- Surveillance et audit
- S'assurer que la chaîne logistique tierce est sécurisée

Système d'exploitation et Kubernetes : misez également sur la réduction de la surface d'exposition aux attaques, avec notamment :

- Analyse des vulnérabilités
- Application régulière de correctifs
- Mise à jour des composants Kubernetes
- Contrôle d'accès basé sur la gestion des identités, accès basé sur les rôles, et accès du moindre privilège
- Sécurisation du plan de contrôle, y compris le serveur API, les codes secrets, le Kubelet et d'autres composants
- Utilisation d'espaces de noms

Applications d'IA générative : mettez en œuvre des mesures de sécurité ciblant les nouvelles surfaces d'exposition aux attaques créées par l'IA générative :

- Gestion des identités pour lutter contre l'injection d'invites, la divulgation de données sensibles, le vol de modèles et la corruption des données d'apprentissage
- Validation des sources de données pour se protéger contre la corruption des données d'apprentissage, biais de modèle
- Surveillance et audit pour identifier et prévenir l'utilisation du DOS sur les modèles, le vol de modèles, la divulgation de données sensibles, détection d'anomalies, analyses approfondies

Données : intégrez des mesures de protection des données solides pour sécuriser les données dans le modèle de langage et l'application :

- Cyber-coffre-fort isolé
- Chiffrement
- Plan de réponse aux incidents
- Surveillance et audit des données et des sorties d'apprentissage

Veillez à ce que les principes de protection des données soient appliqués à toutes les données, y compris les entrées d'apprentissage, les sorties de modèle et toutes les données impliquées dans la Retrieval Augmented Generation (RAG), le cas échéant. En outre, assurez la conformité continue avec toutes les réglementations applicables en matière de protection des données.

Utiliser les principes Zero-Trust

Le rôle de plusieurs principes Zero-Trust tels que la gestion des identités, l'accès avec le moindre privilège, le renforcement du système et l'application de correctifs a déjà été mentionné, ce qui montre la valeur des principes Zero-Trust dans la sécurisation d'une charge applicative de l'IA générative. Les architectures Zero-Trust nécessitent également une journalisation, une surveillance et un audit continus de l'activité du réseau, ce qui peut empêcher les risques spécifiques à l'IA générative tels que la manipulation des résultats et la corruption des données.

DELL Technologies

De plus, le Zero-Trust encourage la micro-segmentation, ce qui réduit l'impact des violations. Il requiert également le chiffrement des données en transit et au repos, contribuant ainsi à la stratégie globale de protection des données.

Outre ces exemples d'application du Zero-Trust pour sécuriser une charge applicative de l'IA générative, adopter les principes du Zero-Trust reste définitivement une bonne pratique en matière de sécurité.

Assurer la gouvernance et la supervision humaine

La valeur de l'IA générative réside en grande partie dans l'automatisation des tâches qu'un humain exécuterait normalement, mais la gouvernance humaine est essentielle pour garantir la sécurité et le bon fonctionnement des applications. Un modèle de gouvernance implique généralement des parties prenantes clés dans l'ensemble de l'entreprise, qui définissent des directives et des exigences en matière de conformité à l'éthique et aux normes, des politiques et procédures de gestion des données, et, enfin, décrit comment sont réparties les responsabilités.

Une gouvernance et une surveillance appropriées peuvent aider à résoudre des problèmes tels que la dépendance excessive aux modèles, les préjugés, la manipulation des résultats, la divulgation d'informations sensibles et la corruption des données.

M. Lederman souligne également l'importance de la formation : « Les utilisateurs doivent être formés. Ils doivent apprendre à utiliser le système d'IA générative, ce qu'il faut faire, mais aussi ce qu'il ne faut pas faire. »

En plus du risque que représentent les applications d'IA générative d'une entreprise, il faut également prendre en compte la multiplication des cyberattaques générées par l'IA générative, qui exigent souvent une intervention humaine. Par exemple, les hackers utilisent des deepfakes pour simuler le comportement humain et rendre beaucoup plus efficaces les attaques par phishing en imitant plus précisément le style d'écriture ou de parole d'un humain. La formation continue et la sensibilisation font partie des moyens les plus efficaces de gérer ces risques, pour renforcer encore une fois l'élément humain.

Profiter des outils de sécurité basés sur l'IA générative dès qu'ils sont disponibles

Bien que l'accent soit mis sur la gestion des risques, l'IA générative a également le potentiel de renforcer les efforts en matière de sécurité. Bien que ces capacités soient encore en phase expérimentale, elles présentent néanmoins des avantages significatifs dans trois domaines clés :

- **Contenu :** création d'une règle de sécurité, formation personnalisée, classification des données et création de rapports
- **Prédiction :** des risques et des attaques, suggestion d'actions correctives
- **Connaissances :** interrogation de l'environnement (communication avec le système), analyse approfondie, automatisation

L'intégration de l'IA générative dans les outils de sécurité pourrait optimiser les capacités des équipes de sécurité, réduire les coûts et renforcer la protection. Tirez parti de ces solutions qui ne cessent de se développer.

Innover en toute confiance

Mais surtout, ne laissez pas les risques de sécurité vous empêcher de tirer parti d'une technologie potentiellement révolutionnaire. L'efficacité, l'automatisation, la réduction des coûts, la résolution de problèmes et la créativité font partie des atouts de l'IA générative pour transformer l'entreprise.

Si l'IA générative nécessite des mesures de cybersécurité robustes et parfois novatrices, celles-ci doivent se mettre au service de l'entreprise, et non compromettre ses résultats. La mise en place d'une stratégie de cybersécurité adaptée devrait donner à l'entreprise la confiance nécessaire pour se développer et innover.

Découvrez comment relever certains des défis actuels majeurs en matière de cybersécurité sur **dell.com/cybersecuritymonth**

DELL Technologies